

*Aus dem Adolf-Butenandt-Institut der  
Ludwig-Maximilians-Universität München  
Lehrstuhl: Molekularbiologie  
Direktor: Prof. Dr. Peter B. Becker  
Arbeitsgruppe: Prof. Dr. Gunnar Schotta*

# **The role of repressive chromatin functions during haematopoiesis**



*Dissertation zum Erwerb des Doktorgrades der  
Naturwissenschaften (Dr. rer. nat.) an der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität München  
vorgelegt von*

**Alessandra Pasquarella**

*aus Potenza*

2014



**Gedruckt mit Genehmigung der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität München**

**Betreuer:** Prof. Dr. Gunnar Schotta

**Zweitgutachter:** Prof. Dr. Rainer Haas

**Dekan:** Prof. Dr. med. Dr. h. c. M. Reiser, FACR, FRCR

**Tag der mündlichen Prüfung:** 16.04.2015





## Eidesstattliche Versicherung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Thema

**“The role of repressive chromatin functions during haematopoiesis”**

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

---

Ort, Datum

---

Unterschrift Alessandra Pasquarella

My doctoral work focused on the description and analysis of the role of repressive histone modifications during haematopoiesis. This thesis includes the three main projects which I followed during my PhD. Although two of them still require further investigation to be suitable for publication, the results described in the section “SETDB1-mediated silencing of *MuLVs* is essential for *B* cell development” have already been assembled in a manuscript which is ready for submission.

# Table of contents

<b>I. ABSTRACT .....</b>	<b>1</b>
<b>II. ZUSAMMENFASSUNG .....</b>	<b>3</b>
<b>1 INTRODUCTION .....</b>	<b>6</b>
<b>1.1 Epigenetic modifications: several ways to regulate gene expression .....</b>	<b>6</b>
1.1.1 DNA methylation .....	6
1.1.2 Nucleosome positioning and histone modifications .....	8
1.1.3 Histone modifications during development .....	10
<b>1.2 Haematopoiesis: a good tool to study chromatin functions .....</b>	<b>14</b>
1.2.1 Central haematopoiesis .....	15
1.2.2 Peripheral haematopoiesis .....	24
<b>1.3 The histone methyltransferase SETDB1 .....</b>	<b>31</b>
<b>1.4 The histone methyltransferases SUV420H .....</b>	<b>33</b>
<b>2 RESULTS (I) .....</b>	<b>35</b>
<b>2.1 Spontaneous germinal center activation in <i>Suv420h2</i> knockout (ko) mice .....</b>	<b>35</b>
2.1.1 <i>Suv420h2 ko</i> mice show enlarged spleen .....	35
2.1.2 <i>Suv420h2 ko</i> B cells spontaneously initiated the germinal center reaction .....	35
2.1.3 <i>Suv420h2 ko</i> B cells response to external stimuli is comparable to wild type cells .....	40
2.1.4 Follicular helper T cells are overrepresented in <i>Suv420h2 ko</i> spleen .....	41
2.1.5 Release of apoptotic checkpoints exacerbates germinal center activation in <i>Suv420h2 ko</i> mice .....	43
2.1.6 <i>Suv420h2 ko; VavBcl2</i> mice have increased proportion of follicular helper T cells .....	45
<b>2.2 Discussion (I) .....</b>	<b>47</b>
<b>3 RESULTS (II) .....</b>	<b>51</b>
<b>3.1 SETDB1-mediated silencing of MuLVs is essential for B cell development .....</b>	<b>51</b>

- 3.1.1 Loss of *Setdb1* impairs B cell development..... 51
  - 3.1.2 Intrinsic role of *Setdb1* during B cell development ..... 56
  - 3.1.3 *Setdb1*<sup>Mbl</sup> B cells ectopically transcribe murine retrotransposable elements..... 58
  - 3.1.4 SETDB1 represses retrotransposons in pro B cells ..... 62
  - 3.1.5 MuLVs derepression results in activation of neighbouring genes..... 64
  - 3.1.6 *Setdb1* loss in pro B cells induces apoptosis..... 65
  - 3.1.7 Overexpression of antiapoptotic *Bcl2* partially rescues B cell defects in *Setdb1*<sup>Mbl</sup> mice 66
- 3.2 Discussion (II) .....69
- 4 RESULTS (III)..... 74
- 4.1 The role of the histone methyltransferase SETDB1 during haematopoiesis.....74
- 4.1.1 *Setdb1*<sup>Vav</sup> mice are underdeveloped and show impaired hematopoietic organs..... 74
  - 4.1.2 Lack of *Setdb1* during hematopoietic development results in complete loss of lymphocytes and expansion of myeloid-erythroid lineage in the periphery ..... 76
  - 4.1.3 *Setdb1*<sup>Vav</sup> bone marrow completely abolishes B cell development in favour of myeloid-erythroid expansion..... 76
  - 4.1.4 Common lymphoid progenitors (CLPs) are not altered in *Setdb1*<sup>Vav</sup> mice while granulocytes macrophage progenitors (GMPs) are overrepresented..... 78
  - 4.1.5 HSCs compartment collapses 4 weeks after birth in *Setdb1*<sup>Vav</sup> mice..... 79
- 4.2 Discussion (III).....81
- 5 APPENDIX ..... 86
- 5.1 Appendix (Results I) .....86
- 5.2 Appendix (Results II).....88
- 6 MATERIAL AND METHODS..... 112
- 6.1 Materials.....112
- 6.1.1 Mice ..... 112
  - 6.1.2 Cell lines ..... 112
  - 6.1.3 Antibodies and dyes..... 112
  - 6.1.4 Technical devices and material..... 113
  - 6.1.5 Kits..... 114

6.1.6	Tissue culture media, cytokines and immunostimulants .....	115
6.1.7	Reagents and buffers .....	115
6.1.8	Oligonucleotides .....	117
6.1.9	Software and databases .....	118
<b>6.2</b>	<b>Methods.....</b>	<b>119</b>
6.2.1	Mice and cell lines .....	119
6.2.2	Flow cytometry and cell sorting.....	119
6.2.3	Red blood cell lysis .....	119
6.2.4	Definition of hematopoietic cell types for FACS analysis and FACS sorting.....	120
6.2.5	Envelope protein and Fcγr2b staining.....	121
6.2.6	Bone marrow transplantation .....	121
6.2.7	Immunohistochemistry.....	121
6.2.8	Autoantibodies detection test on MEFs .....	122
6.2.9	ELISA .....	122
6.2.10	B cell proliferation assay.....	123
6.2.11	Colony forming assay and B cell differentiation on OP9 cells.....	123
6.2.12	Annexin V staining .....	123
6.2.13	Cytospin .....	124
6.2.14	RNA-Seq and qRT-PCR .....	124
6.2.15	Microarray analysis.....	124
6.2.16	Chromatin immunoprecipitation and ChIP-Seq.....	124
<b>7</b>	<b>ABBREVIATIONS .....</b>	<b>127</b>
<b>8</b>	<b>CURRICULUM VITAE .....</b>	<b>131</b>
<b>9</b>	<b>ACKNOWLEDGMENTS .....</b>	<b>132</b>
<b>10</b>	<b>CITATIONS .....</b>	<b>133</b>



# I. Abstract

Histone modifications represent one of the main layers which determine gene regulation. According to the type of modification we can discriminate repressive histone marks, which correlate with gene silencing; from active histone marks, which promote gene activation.

Beyond H3K27me3, which is mostly responsible for silencing of developmental genes, H3K9me3 and H4K20me3 marks are also associated with repressive functions; however they mainly accumulate at constitutively silenced chromatin regions (heterochromatin). Noteworthy, H3K9me3 is also found outside heterochromatin. Here, SETDB1 is the histone methyltransferase responsible for its establishment.

To investigate the role of H4K20me3 repressive mark we used *Suv420h2* knockout mice; while H3K9me3 functions were examined in the context of B cell development and during haematopoiesis by generating *Setdb1*<sup>Mbl</sup> and *Setdb1*<sup>Vav</sup> conditional knockout mice.

*Suv420h2* knockout mice showed increased spleen size. Immunophenotyping of the main splenic populations highlighted the presence of germinal center B cells which produced substantial amounts of self-reactive antibodies. Additionally, we could also detect a 2 fold increase in follicular helper T cells, which are known to initiate the germinal center reaction. Generation of *Suv420h2* knockout; *VavBcl2* mice, which bypass apoptosis via *Bcl2* overexpression, allowed us to verify whether we could enhance the spontaneous germinal center formation by releasing apoptotic restraints. Importantly, these animals robustly confirmed, with better resolution, the phenotype observed in *Suv420h2* deficient mice.

Altogether, these data suggested that loss of SUV420H2-mediated H4K20me3 promoted spontaneous germinal center formation in the spleen.

*Setdb1*<sup>Mbl</sup> mice exhibited impaired B cell development as we observed enhanced apoptosis during pro B to pre B cell transition. Bone marrow transplantation, together with *in vitro* assay confirmed that this defect is cell-autonomous; therefore *Setdb1* is intrinsically required during B cell development. RNA-Seq analysis of pro B cells revealed massive upregulation of two specific classes of retroelements, known as MuLVs and MMLV30, upon *Setdb1* depletion. Interestingly, ChIP-Seq analysis revealed that SETDB1 and H3K9me3 both localized across these endogenous retroviral sequences, indicating that SETDB1 is required to silence these elements in pro B cells. Since ectopic expression of retroviral elements is known to cause apoptosis, we generated *Setdb1*<sup>Mbl</sup>; *VavBcl2* mice. In these animals, pro B cells overcome programmed cell death and, notably, B cell development is partially rescued as demonstrated by the presence of mature B cells in the spleen.

## Abstract

---

In conclusion, we propose that ectopic expression of endogenous retrovirus due to *Setdb1* loss blocks B cell development.

*Setdb1*<sup>Vav</sup> mice were severely underdeveloped and died within 4-5 weeks after birth. Analysis of hematopoietic cells in the spleen and bone marrow highlighted the complete absence of mature lymphocytes in contrast to an abnormal expansion of the myeloid cells compartment. Common lymphoid progenitors were virtually present in the absence of *Setdb1*, however they did not progress further to generate mature B and T cells. Differently, common myeloid progenitors showed a clear bias towards granulocytes macrophage differentiation. Additionally, we found that the hematopoietic stem cells were detectable only till the 3<sup>rd</sup>-4<sup>th</sup> week after birth; in fact, they vanished approximately before *Setdb1*<sup>Vav</sup> mice died.

Preliminary experiments performed *in vitro* suggested that all these hematopoietic defects are cell-intrinsic; however bone marrow transplantations will confirm this assumption with better confidence.

Altogether these data show that *Setdb1* depletion (1) impairs lymphopoiesis downstream common lymphoid progenitors, (2) induces a bias toward myeloid differentiation and (3) prevents hematopoietic stem cells maintenance; suggesting a potential role of *Setdb1* in lineage commitment and hematopoietic stem cell maintenance.



## II. Zusammenfassung

Histonmodifikationen sind eine Form der Genregulation. Es gibt repressive und aktivierende Modifikationen; die Ersteren sind mit Gensuppression und die Zweiten mit gesteigerter Genexpression assoziiert.

Während die H3K27 Trimethylierung für die Suppression von Genen verantwortlich ist, die in der Embryonalentwicklung eine Rolle spielen, haben die H3K9- und H4K20 Trimethylierung suppressive Eigenschaften in anderen Bereichen des Genoms. Diese beiden Modifikationen sind meist in konstitutiv supprimiertem Chromatin (sog. Heterochromatin) zu finden. Zusätzlich ist eine H3K9 Trimethylierung auch außerhalb von Heterochromatin anzutreffen. Dort ist die Methyltransferase SETDB1 für die Etablierung dieser Modifikation verantwortlich.

Um die Rolle der suppressiven H4K20 Trimethylierung untersuchen zu können, wurden *Suv420h2* Knockout-Mäuse erzeugt. Die Funktion der H3K9 Trimethylierung wurde stattdessen in *Setdb1<sup>Mb1</sup>* und *Setdb1<sup>Vav</sup>* Mäusen beleuchtet, da diese eine Charakterisierung während der B-Zell-Entwicklung und der Hämatopoese erlauben.

*Suv420h2* Knockout-Mäuse zeigen eine vergrößerte Milz. Im Rahmen einer Immunophänotypisierung der Haupt-Milzzellpopulationen konnten Keimzentrums-B-Zellen identifiziert werden, welche große Mengen an selbstreaktiven Antikörpern produzierten. Des Weiteren konnte auch eine zweifache Vermehrung der follikulären T-Helferzell-Population detektiert werden, welche bekanntermaßen die Keimzentrumsreaktion initiiert. *Suv420h2* Knockout- *VavBcl2* Mäuse, welche die Apoptose durch überexprimiertes *Bcl2* hemmen, wurden dahingehend überprüft, ob durch die Apoptosehemmung die Rate der spontanen Keimzentrumsreaktion erhöht werden kann. Es bestätigte sich der aus den *Suv420h2* defizienten Mäusen bekannte Phänotyp in diesen Tieren sehr klar und deutlich. Daraus lässt sich schließen, dass der Verlust der von SUV420H2 vermittelten H4K20 Trimethylierung die spontane Keimzentrumsreaktionsrate in der Milz erhöht.

Die phänotypische Analyse der *Setdb1<sup>Mb1</sup>* Mäuse ergab eine gestörte B-Zellentwicklung welche sich durch eine erhöhte Apoptoserate während der pro-B – zu prä-B-Zellreifung auszeichnet. Knochenmarktransplantations- und *in vitro* Experimente konnten zeigen, dass dieser Defekt spezifisch für die B-Zellpopulation ist. *Setdb1* stellt sich somit als essenzieller Faktor für die B-Zellreifung dar. Die Auswertung von RNA-Sequenzierungsdaten einer *Setdb1* Knockout pro-B-Zellpopulation ergab eine signifikante Expressionserhöhung von MuLV- und MMLV30 retroviralen Elementen. Des Weiteren zeigte die ChIP-Seq-Analyse die Kolokalisation von SETDB1 mit H3K9 Trimethylierung in diesen endogenen retroviralen

## Zusammenfassung

---

Sequenzen an, was darauf hindeutet, dass SETDB1 für die Stilllegung dieser Elemente in pro-B-Zellen verantwortlich ist. Da die ektopische Überexpression retroviraler Elemente bekanntlich Apoptose induziert, wurden *Setdb1*<sup>Mb1</sup>; *VavBcl2* Mäuse gezüchtet, in welchen pro-B-Zellen den programmierten Zelltod überstehen können. Bemerkenswerterweise konnten diese Tiere die B-Zellreifung teilweise wieder durchführen, was sich durch die Anwesenheit von reifen B-Zellen in der Milz zeigte. Daraus lässt sich schließen, dass eine durch den *Setdb1* Verlust induzierte ektopische Expression von endogenen retroviralen Elementen die B-Zellentwicklung blockiert.

*Setdb1*<sup>Vav</sup> Mäuse waren physisch deutlich unterentwickelt und starben innerhalb von 4-5 Wochen nach der Geburt. Die Analyse der hämatopoetischen Zellpopulationen in der Milz und im Knochenmark zeigte die völlige Abwesenheit von reifen Lymphozyten und eine auffällige Erhöhung der myeloiden Zellzahl an. In Abwesenheit von *Setdb1* waren myeloide Vorläuferzellen zu finden, die sich allerdings nicht zu reifen B- und T-Zellen entwickelten. Stattdessen differenzierten diese myeloide Vorläuferzellen vermehrt zu Granulozyten und Makrophagen. Zusätzlich konnten hämatopoetische Stammzellen nur bis zur 3.-4. Woche nach Geburt detektiert werden. Diese Zellpopulation verschwindet kurz vor dem Ableben der *Setdb1*<sup>Vav</sup> Mäuse.

Vorläufige *in vitro* Daten deuten an, dass all diese Phänotypen auf zellintrinsische Defekte zurückzuführen sind. Mittels Knochenmark Transplantationsexperimenten kann dies genauer untersucht werden.

Zusammenfassend zeigen diese Daten, dass die *Setdb1* Depletion (1) die Lymphopoese unterhalb der Ebene der lymphoiden Vorläuferzellen stört, (2) die myeloide Differenzierung präferenziell fördert und (3) die Selbstvermehrung von hämatopoetischen Stammzellen verhindert. Diese Phänotypen legen nahe, dass *Setdb1* eine wichtige Rolle während der Determination der Zellliniendifferenzierung und Stammzell-Selbstvermehrung spielt.



# 1 Introduction

## 1.1 Epigenetic modifications: several ways to regulate gene expression

The semantic of the word epigenetics can be directly deduced from the etymologic analysis of the term. The Greek prefix epi-, which precedes the word genetics, indicates something which is beyond, outside of; thus the expression epigenetics means beyond genetics.

In the field of biology, such term was first used by Conrad Waddington (1905–1975), who is given credit for coining the term, defining epigenetics as “the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being”. With this description, Waddington implied that although different organisms have almost identical DNA sequence, there must be something else, beyond genetics, which determines their different phenotypes (Dupont et al., 2009; Goldberg et al., 2007; Waddington, 1942).

As time passes, this term evolved and developed, encompassing many other meanings as result of the big progresses which have been achieved during the past years in the field of molecular biology. The most recent and comprehensive meaning of the word epigenetic was discussed during a meeting (December 2008) hosted by the Banbury Conference Center and Cold Spring Harbor Laboratory. Here, a group of experts attempted to deliver a consensus definition of “epigenetics” which formalized in the following statement: “An epigenetic trait is a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence” (Berger et al., 2009). Although this definition is understandable and broadly accepted from people of the field, biologists with different expertise and background might have some difficulties to comprehend how changes in gene expression, which are possibly inheritable, occur without any modification of the DNA sequence. To circumvent these difficulties, it would be appropriate, at this point, to explain more in detail how epigenetics modifications occur and what they are.

According to the molecular machinery which mediates them, epigenetic modifications can be classified in three main categories: DNA methylation, nucleosome positioning and histone modifications (Portela and Esteller, 2010).

### 1.1.1 DNA methylation

The chemical deposition of methyl groups on DNA is a process known as DNA methylation. DNA methylation mainly occurs on the cytosine residues belonging to cytosine/guanine pairs (CpGs), where nucleotides are separated by only one phosphate group. In mammals, 60-90% of CpG dinucleotides are methylated (Kass et al., 1997) .

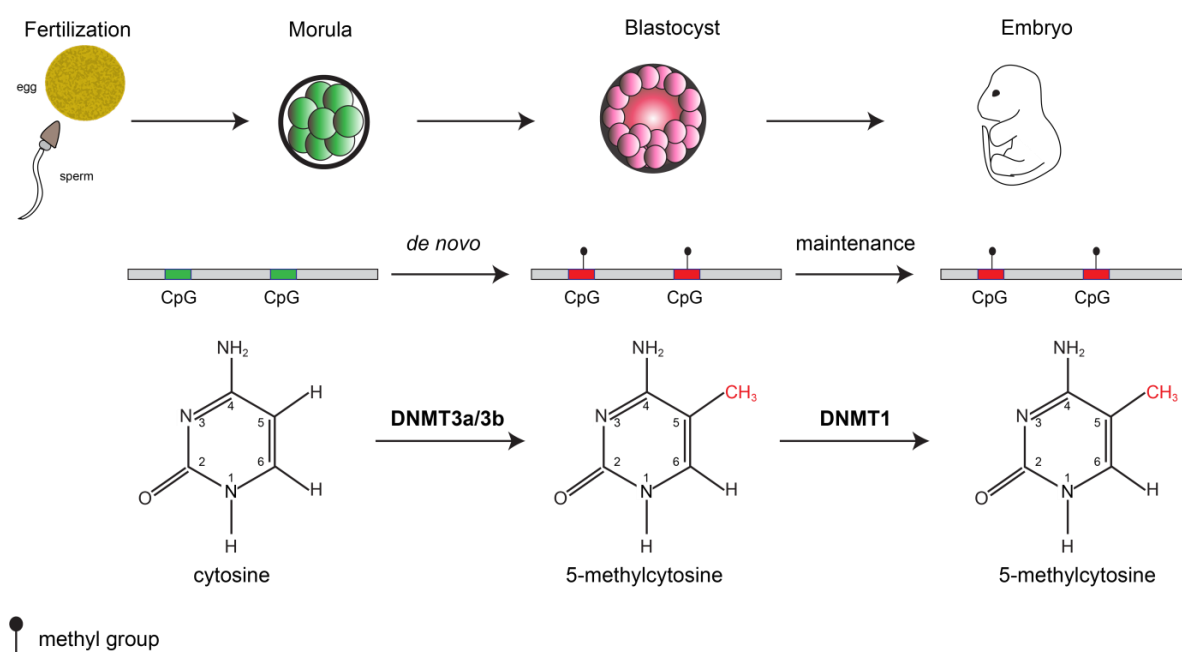
## Introduction

DNA methylation is mainly associated with transcriptional repression and it plays an essential role in important biological processes, such as genomic imprinting, X-chromosome inactivation, suppression of repetitive elements and gene silencing (Newell-Price et al., 2000; Robertson, 2005; Straussman et al., 2009).

The enzymes which catalyse the transfer of the methyl group on the DNA belong to the family of DNA methyltransferases (DNMTs). In mammals, five DNMTs have been described: DNMT1, DNMT2, DNMT3A, DNMT3B DNMT3L; however among them only DNMT1, DNMT3A and DNMT3B possess methyltransferase activity (Bestor, 2000).

DNMT3A and DNMT3B are *de novo* DNA-methyltransferases responsible for the establishment of a new methylation pattern during early stages of embryo development, when paternal and maternal DNA are deprived of all pre-existing methylation profiles (Okano et al., 1999).

Once that the new methylation profile has taken shape, it needs to be kept each time cells replicate their DNA during the cell cycle (S phase). At this stage the newly synthesized genome consists of a hybrid DNA molecule which contains an old methylated strand, and a new one without methyl groups. To restore the methylation pattern on the newly synthesized strand, the methyltransferase DNMT1 is recruited at replication foci to maintain the methylation pattern previously established by DNMT3A and DNMT3B (Bostick et al., 2007; Chuang et al., 1997) (**Fig. 1.1**).



**Figure 1.1 DNA methylation**

Upon fertilization maternal and paternal DNA are demethylated. In the blastocysts DNMT3A and DNMT3B will establish a new methylation profile which will be kept in the organism by DNMT1.

**1.1.2 Nucleosome positioning and histone modifications**

Chromatin is the way eukaryotic DNA and protein are packed inside cell nuclei. The essential units which give shape to the chromatin are the nucleosomes; structural blocks consisting of an octamer containing two copies of the 4 main histones (H3, H4, H2A and H2B) around which 147 base pair of DNA are enfolded (Luger et al., 1997). All histones are nuclear proteins characterized by a spherical shape which harbour N-terminal tails carrying different kinds of covalent modifications on different amino acid residues.

The fact that DNA is highly packed around histones has a strong impact on the transcriptional level. Nucleosomes, in fact, represent a physical barrier which prevents access of transcription factors to their binding sites in the genome (Jiang and Pugh, 2009). Depending on the position of nucleosomes along genes, transcriptional blockage can occur at different moments during transcription. Indeed, nucleosomes positioned at the transcriptional start site (TSS) of a gene prevent the binding of protein complexes involved in the initiation of the transcriptional event (Cairns, 2009); while nucleosomes located in the gene body or distant from the TSS will inhibit later stages of gene expression (transcriptional elongation) (Subtil-Rodriguez and Reyes, 2010). When transcription begins, nucleosomes are relocated along DNA sequence, allowing the binding of the transcriptional machinery (Jiang and Pugh, 2009; Segal and Widom, 2009; Struhl and Segal, 2013) (**Fig. 1.2 a**).

Nucleosome mediated gene regulation has been extensively studied in yeast, and many progresses have been made in understating the molecular basis of this phenomenon and which chromatin remodelers are involved (Segal and Widom, 2009; Struhl and Segal, 2013). However, whether this mechanism occurs also in mammal still remains an open question.

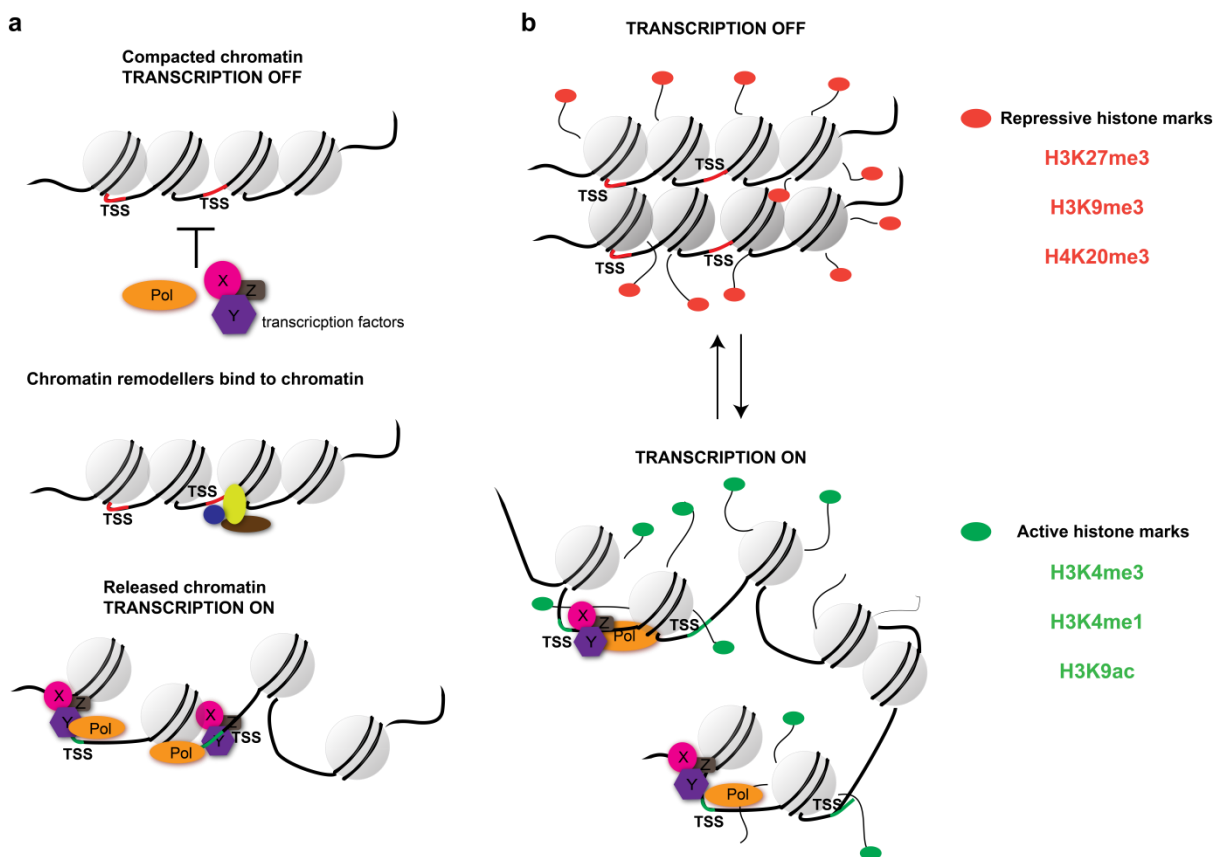
Differently from nucleosome positioning, histone modifications and the impact they have on gene expression have been studied in a broader spectrum of organisms, mammals included. Histone tails undergo post-transcriptional modifications on specific amino acid residues, where several types of functional groups are covalently bound. In the past decades, many histone modifications, such as methylation, acetylation, ADP-ribosylation, ubiquitination, citrullination and phosphorylation have been identified (Kouzarides, 2007). Among these, methylation and acetylation are better described and have been confirmed to alter DNA-protein interaction, influencing transcription (Dupont et al., 2009).

During the last years many ChIP-Seq experiments have been performed to map histone modifications genome-wide (Schones and Zhao, 2008). Interestingly, their distribution

## Introduction

strongly correlates with specific transcriptional states, supporting the assumption that histone marks play a crucial role in regulating genes (Strahl and Allis, 2000).

Chromatin can be distinguished in transcriptionally repressed chromatin and transcriptionally active euchromatin. Repressed chromatin is marked by the presence of methylated lysine residues that associate with transcriptional inactivation (H3K9me3, H3K27me3, and H4K20me3). Euchromatin, on the other side, is enriched either with methylated (H3K4me3, H3K36me3, and H3K79me3) or acetylated (H3K9ac, H3K14ac, H3K18ac, H3K23ac, H4K5ac, H4K8ac, H4K12ac, and H4K16ac) lysine residues which correlate with gene expression (Dupont et al., 2009).



**Figure 1.2 Nucleosome positioning and histone modifications influence transcription**

**a)** Nucleosome positioned at TSS. Transcription is OFF. Upon binding of chromatin remodelers nucleosomes move elsewhere and free TSSs allowing binding of the transcriptional machinery.

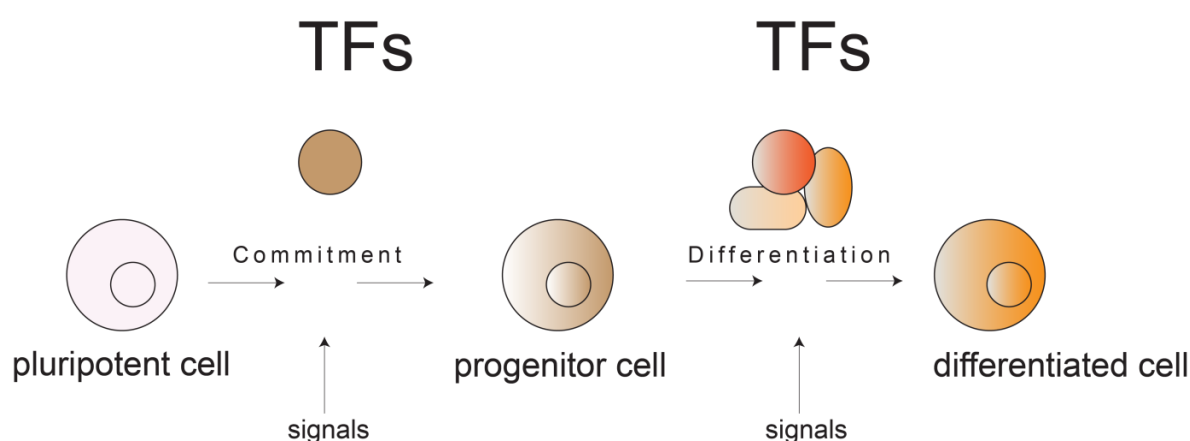
**b)** Repressive histone marks package chromatin in a close configuration preventing transcription. Active histone marks, instead, release compacted chromatin giving access to transcription factors to initiate transcription.

The establishment of these histone modifications all over the DNA sequence is a process anything but trivial as it involves the assembly of big protein complexes, such as MLL (H3K4me3) or the Polycomb protein complex (H3K7me3) in response to certain stimuli. The

group-transferase activity is attributed only to the enzyme which is recruited in such complexes and is essential for the establishment of histone tails modifications. However; the presence of other factors not directly involved in histone tail modification is also crucial for complex assembly, reaction dynamics and sequence targeting in a context dependent manner (Glass and Rosenfeld, 2000; Rosenfeld et al., 2006; Smith and Shilatifard, 2010). This has been largely demonstrated in mice or cell lines by loss of function experiments. In fact, the depletion of non-enzymatic factors which are part of these epigenetic complexes is sufficient to impair the establishment of histone modifications (Ang et al., 2011; Dou et al., 2006; Faust et al., 1995; Katada and Sassone-Corsi, 2010; Pasini et al., 2004; Rampalli et al., 2007). Significantly, loss or reduction of histone marks induce changes in transcription (Kouzarides, 2007), supporting the hypothesis that the histone modifications represent one of the main layers which regulates gene expression (**Fig. 1.2 b**).

### 1.1.3 Histone modifications during development

Because histone modifications play a pivotal role in fine-tuning gene expression, they are often investigated in the context of development and differentiation (Dambacher et al., 2010). These two biological processes involve massive transcriptional changes driven by specific transcription factors. Up to the present, several publications have already revealed a lot about the importance of transcription factors during developmental transitions (Dambacher et al., 2010), but we still lack some knowledge about their interplay with the chromatin environment (**Fig. 1.3**).



**Figure 1.3 Transcription factors (TF) determine cell commitment**

Transcription factors set out transcriptional networks which promote developmental switches and terminal differentiation.



## Introduction

---

Developmental switches as well as cell differentiation require activation of many lineage specific genes; therefore at many sites of the genome active histone modifications have to be installed to generate a permissive chromatin environment. So far, the best investigated histone methyltransferases associated with gene activation are MLL1-5 (Mixed-Lineage-Leukaemia), the Thrithorax homologs in *Drosophila* (Yu et al., 1998). MLLs mediate H3K4me3 establishment via its C-terminal SET domain, promoting gene expression (Krivtsov and Armstrong, 2007; Milne et al., 2002). Similarly to other histone methyltransferase, MLLs cooperates in a big complex with many other proteins, such as WDR5, RbBP5 and ASH2L, which support the enzymatic activity (Dou et al., 2006). Likewise Trithorax complex in *Drosophila*, MLL1-5 target important developmental factors in vertebrates. In particular, loss of MLLs prevents activation the homeotic genes, a group of transcription factors traditionally used as model to study developmental gene regulation (Milne et al., 2002; Yu et al., 1995).

In order to study the developmental function of distinct MLL family members, several knockout mice have been generated. *Mill* knockout (ko) mice die during embryogenesis, between E12.5 and E16.5, probably due to abortive expression of some homeotic genes (Yu et al., 1998; Yu et al., 1995). Remarkably, disruption of only one MLL1 allele (*Mill* heterozygous mice) is already sufficient to induce several defects as retarded growth, haematopoietic abnormalities, defects in the axial skeleton formation and sternal malformations (Yu et al., 1995). To verify whether the observed developmental defects were linked to MLL1 methyltransferase activity, mice carrying the *Mill* gene depleted of the SET domain (*Mill*ΔSET) have been also investigated. Surprisingly, these animals are viable and fertile; demonstrating that the role of MLL1 during development is only partially linked to its enzymatic activity (Terranova et al., 2006). However, expression of some homeotic genes was reduced in *Mill*ΔSET mice, indicating that MLL's gene activation function is directly linked to the enzymatic properties of MLL1 (Terranova et al., 2006). Developmental defects during embryogenesis have also been observed upon disruption of *Mill2* gene. *Mill2* ko mice showed embryonic lethality at E11.5 due to growth retardation, apoptosis and developmental delay. However, this phenotype is milder compared to MLL1 loss of function. Although *Mill* and *Mill2* are both essential during embryogenesis, their depletion results in different phenotypes and misregulation of different sets of homeotic genes; indicating that these two factors control independent aspects of embryogenesis (Glaser et al., 2006).

Notably, *Mill* ko embryos fail in expanding the hematopoietic stem cell compartment and loss of only one *Mill* allele also results in hematopoietic defects, supporting the idea that *Mill* has a defined role in this context (Yu et al., 1995). Unfortunately, embryonic lethality blocked any exploratory attempt to elucidate function of MLL1 in adult tissue. For this

purpose, conditional deletion has been adopted to study *Mlll* functions in the context of adult development, specifically during haematopoiesis. Tissue-specific deletion of *Mlll* was achieved in two independent studies, using different strategies (Jude et al., 2007; McMahon et al., 2007). Noteworthy, the usage of different hematopoietic specific cre recombinases resulted in inconsistent phenotypes. However, in both works, the observation that *Mlll* is important for HSCs functions and self-renewal was a common denominator (Jude et al., 2007; McMahon et al., 2007), corroborating the hematopoietic defects observed in the embryos.

A broad range of hematopoietic anomalies have been also observed upon loss of *Mll5* (Liu et al., 2009). Although *Mll5 ko* mice showed some postnatal lethality and were clearly distinguishable by reduced size, surviving animals evolved as healthy adults (Madan et al., 2009). Remarkably, hematopoietic anomalies were observed in three independent publications where different strains of *Mll5 ko* mice were generated via comparable knockout approaches. Intriguingly, each of the knockout strain exhibited a big variety of hematopoietic defects, spanning from impaired long term repopulation capabilities of hematopoietic stem cells to eye infection caused by impaired neutrophil functionality (Heuser et al., 2009; Madan et al., 2009; Zhang et al., 2009). Outside the hematopoietic context, other investigators underlined the importance of MLL3-4 mediated gene activation during fat generation. In fact, it seems that these two proteins are included in the so called ASCOM complex, which mediates PPAR $\gamma$ -dependent adipogenesis (Lee et al., 2008).

Although the importance of active histone marks is unquestionable, cell commitment also requires silencing of lineage inappropriate genes/transcripts. Thus, formation of repressive chromatin is also necessary to ensure proper differentiation (Dambacher et al., 2010).

Transcriptionally inactive regions are clearly distinguishable by the presence of H3K27me3 and H3K9me3. Even though these histone modifications both correlate with repressive feature, they mutually exclude each other at transcriptionally inactive chromatin. This has been demonstrated in different cell types by genome-wide studies which consistently confirmed that H3K9me3 and H2K27me3 co-occupancy is hardly found throughout the genome (Kim and Kim, 2012; McEwen and Ferguson-Smith, 2010). Consequently, it is reasonable to assume that H3K9me3 and H2K27me3 associate with completely different sets of genes and that the protein complexes which establish these marks work independently.

H3K27me3 is mostly associated with promoter of developmental genes (Voigt et al., 2013). At these sites, H3K27me3 establishment is mediated by the Polycomb group proteins (PcGs) repressive complex. These proteins were first identified in *Drosophila melanogaster* as repressors of the homeotic and segmentation genes (Lewis, 1978; Moazed and O'Farrell, 1992;

## Introduction

---

Pelegri and Lehmann, 1994; Ringrose and Paro, 2004). Several PcG proteins are highly conserved in mammals and can be distinct in two classes according to the protein complexes they associate with and to the function they exert (Bracken and Helin, 2009).

The establishment of H3K27me3 is mediated by the PCR2 complex. In mice, the enzymatic activity within this complex is held by the histone methyltransferase enhancer of zeste homologue 2 (EZH2) in immature cells and enhancer of zeste homologue 1 (EZH1) in terminally differentiated cells. These two histone methyltransferase work together with embryonic ectoderm development (EED) and suppressor of zeste 12 homolog (SUZ12), the other two components of the PCR2 complex (Margueron and Reinberg, 2011). On the other end, the PRC1 complex is in charge to recognize and target H3K27me3, propagating it and stabilizing it. Recognition of H3K27me3 marks is mediated by one of the chromodomains contained in the core proteins (PCs) (Fischle et al., 2003; Min et al., 2003). Other members of the PRC1 complex are PSC (Posterior sex combs), PH, SCML (Sex combs on midleg) and RING proteins. In particular, the latter is responsible for the H2A ubiquitination on the lysine 119. This modification prevents the binding of the FACT protein complex, which mediates nucleosomes disassembly (Belotserkovskaya and Reinberg, 2004; Wang et al., 2004).

Although PRC1 and PCR2 complexes have a strong association with H3K27me3, the mechanism which recruits them at the target sites and the way they stabilize chromatin repression are still under debate.

The first attempt to study H3K27me3 during mouse development was made by O'Carroll and colleagues, whom tried to generate at first *Ezh2* deficient mice. Loss of *Ezh2* resulted in embryonic lethality between day 7 and day 8.5. Inability to work with embryonic tissues prompted ESCs generation from *Ezh2 ko* blastocyst to study EZH2 function at the cellular level. However, even this option was negated by the impossibility to derive ES cells from *Ezh2* deficient blastocysts (O'Carroll et al., 2001). More recently, alternative deletion strategies to disrupt *Ezh2* allele allowed the generation of *Ezh2 ko* ESCs, demonstrating a non-essential role of EZH2 in ESC establishment and maintenance (Shen et al., 2008). Curiously, *Ezh2 ko* ES cells did not show activation of those genes where H3K27me3 was lost, probably due to lack of site specific transcription factors. Furthermore, *Ezh2 ko* ESCs kept residual H3K27me3 which was found to be maintained by EZH1, indicating that this enzyme is used as a backup H3K27 histone methyltransferase to protect key developmental genes and to retain pluripotency, preventing unwanted differentiation (Shen et al., 2008).

EZH2 has also been investigated during haematopoiesis. Deletion of *Ezh2* in early B cells development perturbs B cell differentiation; however the mechanisms which leads to this phenotype has been associated to aberrant VDJ recombination rather than altered gene

expression due to loss of H3K27 methyl marks (Ezhkova et al., 2009). *Ezh2* depletion seemed to have no impact on mature B cells development and functionality. Yet, the marked increased expression of the enzyme in the very specialized germinal center B cell population prompted further investigations. Indeed, the supposition that loss of *Ezh2* in mature B cells does not induce any phenotype is valid, but limited to the B cell steady-state. Beguelin *et al.* demonstrated that *Ezh2 cko* B cells are incapable to generate germinal centers, very specialized structures containing highly proliferating antibody producing B cells which are formed upon immune challenge. This deficiency was attributed to loss of H3K27me3 repressive mark on those genes which drive germinal center B cells to terminal differentiation in plasmablasts (Beguelin et al., 2013).

Lastly, developmental mechanisms linked to Polycomb-mediated repression have been elucidated in the context of skin development. Here, EZH2 controls proliferation and prevents premature differentiation of progenitor cells repressing specific genes (Ezhkova et al., 2009).

Beyond H3K27me3, H3K9me3 and H4K20me3 are the other repressive histone marks which accumulate at non-coding regions (centromeres and telomeres) where they are sequentially deposited by the histone methyltransferase SUV39H and SUV420H (Dambacher et al., 2013). Notably; outside these regions, H3K9me3 is also found and the histone methyltransferase are known to take charge of its establishment is SETDB1. However, differently from H3k27me3, the role that H3K9me3 plays during developmental transition is still poorly investigated.

## **1.2 Haematopoiesis: a good tool to study chromatin functions**

Haematopoiesis is the biological process during which multipotent hematopoietic stem cells (HSCs) generate all blood cells, which are classified as lymphoid (T, B and NK) or myeloid-erythroid cells (granulocytes, macrophages, erythrocytes and megakaryocytes). In mouse, HSCs start to be differentiated from the haemogenic endothelium at day 8.5 of embryonic development to relocate afterwards in the foetal liver, where they remain until birth (Clements and Traver, 2013). After birth, the bone marrow gradually becomes the main source of the HSCs (Mikkola and Orkin, 2006). Here, B and myeloid cells accomplish to develop, while T cell progenitors leave this site and travel through the blood stream to complete differentiation in the thymus (Janeway, 2008a). Since bone marrow and thymus support early stages of haematopoietic development, they are defined as primary or central hematopoietic organs. When blood cells have reached a certain degree of maturity at these sites, they leave to reposition in the peripheral or secondary hematopoietic organs (spleen, tonsils, Payer's patches and lymph nodes), where they terminally differentiate in immune effectors.

Central and peripheral haematopoiesis define the spatiotemporal axis along which a lot of differentiation takes place. In fact, they both include highly controlled developmental switches that are coordinated by transcription factors via establishment of new gene expression patterns. Each developmental stage can be identified by surface marker expression, allowing the purification of well-defined hematopoietic subpopulation. The possibility to isolate distinct subpopulations by FACS analysis and the multiplicity of differentiation events, make of the haematopoiesis an ideal system to study development.

### 1.2.1 Central haematopoiesis

In the bone marrow HSCs are retained in specific niches where they are maintained and regulated (Ding and Morrison, 2013; Morrison and Scadden, 2014). Their potential to generate all differentiated blood cells is determined by the exit from the quiescent state which characterizes them. In fact, HSCs retain a low-proliferating profile until the blood system has to be refilled and/or the HSCs pool has to be expanded (Wilson et al., 2008). Each time HSCs proliferate, they asymmetrically divide in self-renewing copy of themselves and multipotent progenitors without self-renewing features (Beckmann et al., 2007; Brummendorf et al., 1998; Takano et al., 2004). The exit from the quiescent state results from a complex interplay between cell intrinsic and cell extrinsic events. On one side, bone marrow microenvironments continuously provide external signals which dictate the transcriptional landscapes that have to be settled during each developmental stage. On the other side, transcription factors represent the cell intrinsic interface which mediates the cross talk between signalling and transcriptional changes.

Loss and gain of function studies broadly investigated the role that different transcription factors play during haematopoiesis (Orkin and Zon, 2008), allowing to outline the hierarchical differentiation model (**Fig. 1.4**). Although this model is constantly challenged by new discoveries, the importance of transcription factors in orchestrating hematopoietic cell commitment is undeniable.

The most striking property of HSCs is to self-renew. In mice, this capability was discovered by long term reconstitution experiments where HSCs have been serially transplanted over a long period of time, without losing their potency (Iscoe and Nawa, 1997; Lemischka et al., 1986). Defective HSCs self-renewal was often linked to inability to maintain quiescence (Li, 2011; Pietras et al., 2011). Loss of the transcription factors GFI1, E2A BMI and LDB1, for example, directly resulted in activation of cell cycle genes. This event released HSCs from quiescence and impaired HSCs long term repopulation capabilities, due to ablation of cell

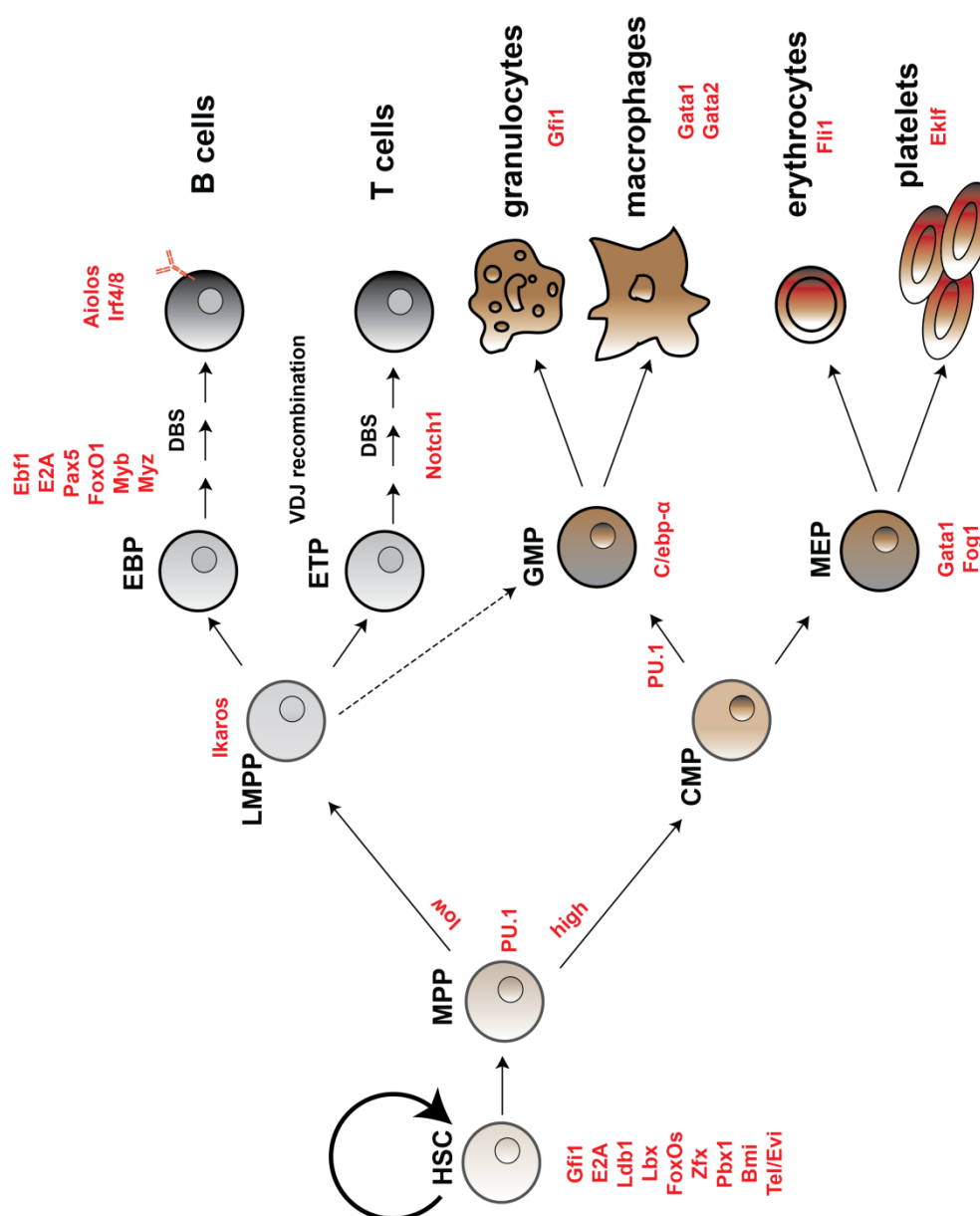
cycle restraints (Hock et al., 2004; Li et al., 2011; Park et al., 2003; Semerad et al., 2009; Zeng et al., 2004). Alteration of proliferation and apoptosis, induction of hypoxia and oxidative stress due to respective depletion of *Evi1*, *FoxOs* and *Zfx* genes also resulted in unwanted HSCs cell cycle entry; however the pathways which led to loss of quiescence were, in these specific cases related to cellular stress rather than direct cell cycle related-genes derepression (Galan-Caridad et al., 2007; Goyama et al., 2008; Tothova et al., 2007).

HSCs quiescence is also controlled by signalling events. Indeed, more recent work demonstrated that HSCs dormancy is perturbed in *Pbx1 ko* mice as a consequence of aberrant expression of genes associated to TGF $\beta$  pathway (Ficara et al., 2008). Additionally, constitutive expression of  $\beta$ -catenin; a key component of Wnt signalling pathway; also caused hematopoietic stem cell failure induced by enforced cell cycle entry (Scheller et al., 2006).

Long term HSCs (LT-HSCs) which lost self-renewal properties become multipotent progenitors (MPPs). During this transition all HSCs gradually turn into short term HSCs (ST-HSCs) with limited self-renewal ability (Adolfsson et al., 2001; Yang et al., 2005). Differently from LT-HSCs, MPPs cannot self-renew but are still multipotent. These cells embody the first hematopoietic branch point where cells start to be hijacked towards myeloid or lymphoid fate. The possibility to isolate very well defined hematopoietic subpopulations via immunophenotyping allowed to investigate *in vivo* and *in vitro* their differentiation potential and to identify the transcription factors which are determinant for lineage choice. LT- and ST-HSCs are included in the Lin<sup>-</sup> Sca<sup>+</sup>, c-kit<sup>+</sup> (LSK) population. Within these cells the expression of FLT-3 (FMS-related tyrosine kinase 3) seems to mark the gradual loss of myeloid potential. LSK FLT3<sup>+</sup> cells are known as lymphoid multipotent primed progenitors (LMPPs) and retain lymphoid-myeloid potential but are incapable to produce megakaryocytes and erythrocytes (MegE) (Adolfsson et al., 2005; Nutt and Kee, 2007; Welner et al., 2008). *Flt3 ko* and *Flt3l ko* mice both significantly reduce production of B cell precursors together with impaired myeloid differentiation (Mackarechtschian et al., 1995; McKenna et al., 2000), indicating that *Flt3* expression is important for lymphoid multipotent primed progenitors specification. Up to know no transcription factors have been identified as definite regulators of the MPPs to LMPPs transition; however there are some beliefs that the ETS (E26 transformation-specific or E-twenty-six) family member PU.1 (*Sfpi1*) might be one of those (Nutt and Kee, 2007). *Sfpi1 ko* mice died during late embryogenesis and showed severe impairment of myeloid and lymphoid lineage, indicating that PU.1 functions are required at very early developmental stages (McKercher et al., 1996; Scott et al., 1994). Artificial introduction of *Sfpi1* in *Sfpi1 ko* foetal liver cells rescued the hematopoietic defects in both lineages. Interestingly, this resulted in differentiation of macrophage which highly expressed

ectopic PU.1 and B cells expressing significant lower amount of the protein, moving the hypothesis that PU.1 dosage is determinant for myeloid versus lymphoid lineage choice (DeKoter and Singh, 2000).

In support of this hypothesis, Rosenbauer and colleagues demonstrated that reduced *Pu.1* expression achieved by deletion of its enhancer upstream regulatory element (URE) caused significant loss of myeloid and conventional B cells (B2) while B1 cells, which are only involved in humoral response, were increasing with aging (Rosenbauer et al., 2006).



**Figure 1.4 Hierarchical model of hematopoietic development**

Hematopoietic stem cells lose their self-renewal properties and start to differentiate giving rise to all blood cells. Each developmental step is defined by the expression of one or more transcription factors which expression is stage dependent.

Surprisingly, conditional deletion of *Sfp1* in committed B cell (Polli et al., 2005; Ye et al., 2005) or *Sfp1* inactivation in sorted CLPs (Iwasaki et al., 2005) did not interfere with the B cell differentiation program, indicating that PU.1 is dispensable for late B cell development.

Besides PU.1, another essential transcription factor which determines lymphoid fate specification is the DNA binding protein IKAROS, encoded by *Ikzf1* gene (Molnar and Georgopoulos, 1994). As previously described, *Ikaros* deficient mice showed severe impairment of lymphopoiesis due to complete abrogation of the B cell compartment and defective T cell development, while the myeloid compartment is unaffected (Georgopoulos et al., 1994; Georgopoulos et al., 1992; Georgopoulos et al., 1997; Wang et al., 1996). Notably, while *Pu.1* deficient mice lack myeloid and lymphoid cells, *Ikaros* loss of function only affects lymphoid specification, suggesting that *Ikaros* comes after *Pu.1* in the lymphoid transcriptional hierarchy (Nutt and Kee, 2007). In the first place, B cell loss in *Ikaros* deficient mice was attributed to lack of LMPPs which were not detectable due to loss of *Flt3* expression (Nutt and Kee, 2007). However, artificial expression of a reporter construct carrying *Ikaros* regulatory elements in wild type and *Ikaros* deficient progenitors suggested that LMPPs are only virtually lost, supporting the idea that *Ikaros* promotes lymphopoiesis downstream LMPPs (Yoshida et al., 2006). Significantly, B cell developmental defects could also be detected in transgenic mice carrying an *Ikaros* hypomorphic allele or an *Ikaros* allele deprived of the DNA binding domain. In these animals, B cells arrested during pro B to pre B and pre B to immature B developmental transition, respectively (Joshi et al., 2014; Kirstetter et al., 2002).

In the light of the most recent discoveries intersecting roles of PU.1 and IKAROS have been proposed to determine lymphoid fate; however the exact mechanisms which modulate the expression of these two transcription factors and the way they convey progenitors to specific fates, are still poorly understood (Nutt and Kee, 2007).

Current models support the idea that PU.1 and IKAROS work in concert to promote FLT3 expression on cell surface (Singh et al., 2005). FLT3 high cells (LMPPs) are considered the earliest lymphoid progenitors (Welner et al., 2008). Indeed, 60% of LMPPs generate nature killer (NKs) cells and 31% start to express lymphoid transcripts (Welner et al., 2008). Although 38% of LMPPs still possess myeloid-lymphoid potential, a percentage of these cells it is already restricted to the lymphoid lineage (Mansson et al., 2007). This subpopulation represent the early lymphoid progenitors (ELPs) which are characterized by loss of VCAM surface marker and individual or synchronized expression of lymphoid specific genes, as recombination activating gene 1 (*Rag1*), interleukin-7 receptor- $\alpha$  (*Il-7ra*) and the terminal



## Introduction

---

deoxynucleotidyl transferase (*Dntt*) (Igarashi et al., 2005; Lai and Kondo, 2006, 2007; Medina et al., 2001). ELPs are the joining link between LMPPs and common lymphoid progenitors (CLP), the next hematopoietic branch point where B, T and NK cell fate is determined *in vivo*; (Allman et al., 2003; Harman et al., 2006; Kondo et al., 1997) but not *in vitro*, as sorted CLPs still show limited myeloid potential (Balciunaite et al., 2005; Nutt and Kee, 2007; Rumfelt et al., 2006; Traver and Akashi, 2004). CLPs are typically defined by surface expression of interleukin-7 receptor- $\alpha$  (IL-7R $\alpha$ ). IL-7R $\alpha$  ligand is the interleukin-7 (IL-7), the cytokine which sustains proliferation and survival of developing lymphocytes (Dias et al., 2005; Kikuchi et al., 2005). Once IL-7R $\alpha$  engages IL-7, a cascade of phosphorylation events takes place, culminating in the recruitment of STAT5A and B. These, are IL-7R $\alpha$  pathway restricted transcription factors which activate the antiapoptotic proteins BCL2, BCL-XL and MCL-1 (Clark et al., 2014; Heltemes-Harris et al., 2011; Kikuchi et al., 2005; Malin et al., 2010). Now on, downstream IL-7R $\alpha$  signalling, a group of well investigated transcription factors start to orchestrate B and T cell development.

E2A, EBF1, FOXO1, MYB; MIZ and PAX5 are the most well characterized transcription factors which coordinate B cell development. Particularly, MYB and MYZ mediate B cell survival and proliferation induced by activation of the IL-7R $\alpha$  pathway (Fahl et al., 2009; Kosan et al., 2010; Thomas et al., 2005); while all the other factors work in concert to carve B cell specific transcriptional networks. CLPs commitment to the B cell fate is signed by the expression of *E2a* and *Ebf1* (Kee, 2009). In mouse, individual depletion of these factors results in similar phenotypes, like early B cell developmental arrest at comparable stages and lack of expression of many B cell-associated genes (Bain et al., 1994; Gyory et al., 2012; Li et al., 1996; Vilagos et al., 2012), indicating that E2A and EBF1 transcription factors have a cognate role in B cell determination (O'Riordan and Grosschedl, 1999; Seet et al., 2004). Notably, developmental arrest shown by *E2a* deficient B cell could be rescued by *Ebf1* overexpression but not vice versa, indicating that E2A works upstream EBF1. Indeed, as demonstrated via ectopic expression in macrophages, E2A specifically binds to *Ebf1* distal promoter ( $\alpha$ -promoter), inducing transcriptional activation (Roessler et al., 2007). EBF1 induces a new wave of gene expression which drags cells to the pro B cell stage passing through the pre-pro B precursors (Lin et al., 2010; Rothenberg, 2010; Zandi et al., 2008). During this transition, two central events take place: the immunoglobulin heavy chains start to rearrange and the B cell master regulator PAX5 is transcribed (Cobaleda et al., 2007; Medvedovic et al., 2011). Although very little is known about the mechanisms which lead to *Pax5* gene activation, genetic analysis designated EBF1 and STAT5 as direct *Pax5* regulator (Hirokawa et al., 2003; O'Riordan and Grosschedl, 1999). Intriguingly, *Pax5* expression is

concomitant with increased level of *Ebfl*. This is due to a feed forward system, where *Ebfl* raises its own expression by directly binding its distal  $\alpha$ -promoter and by activating *Pax5* expression. Indeed, PAX5 boosts *Ebfl* transcription via binding its proximal ( $\beta$ ) promoter, which is different from the one targeted by E2A, suggesting that distinct complexes induce *Ebfl* expression at different developmental stages (Nutt and Kee, 2007; Roessler et al., 2007). The discovery that *E2a* expression is promoted by EBF1-mediated suppression of the inhibitory factor *Id2* and *Id3* (Thal et al., 2009); the finding that *Ebfl* expression is increased and maintained by PAX5 during B cell development (Fuxa et al., 2004) and the fact that PAX5 binding sites often co-localize in the genome with EBF1 (Lin et al., 2010), strongly suggest that B cell commitment is established by a self-reinforcing circuit where several transcription factors teamwork to set out the right transcriptional pattern (Singh et al., 2005).

Although *E2a* and *Ebfl* truly initiate B cell specification and their absence results in loss of B cell specific transcripts; PAX5 is the real guardian angel of B cell development (Cobaleda et al., 2007; Nutt et al., 1999). *Pax5 ko* B cells block at early pro B cells stage (Urbanek et al., 1994); as they exhibited B220 and c-KIT (CD117) markers on their surface (Nutt et al., 1997) and expressed the transcriptional ancestors *E2a* and *Ebfl* together with all their target genes (Cobaleda et al., 2007). The surprising possibility to expand *in vitro Pax5 ko* cells (Nutt et al., 1997) revealed very interesting features of them as, for example, lack of B cell commitment (Cobaleda et al., 2007). In fact, while coculture with stromal cells and IL-7 simply propagated *Pax5 ko* B cells as B220+, c-KIT+ precursors; culture medium implemented with non-lymphoid cell stimulating cytokines induced transdifferentiation into myeloid and NK cells (Nutt et al., 1999). Moreover, *Pax5 ko* B cells could be serially transplanted in recipient mice, denoting recovery of self-renewing features (Rolink et al., 1999).

In wild type cells PAX5 is a transcriptional repressor and activator at the same time (Cobaleda et al., 2007). This double role has been featured by transcriptional analysis which showed downregulation of B cell specific genes and upregulation of lineage inappropriate transcripts in *Pax5 ko* B cells (Nutt et al., 1999; Nutt et al., 1998). These results strongly suggest that commitment does not only involve gene activation, but also requires a tight control over those transcripts which might alter cell fate. For example, ectopic expression of the a PAX5 target gene *Flt3* hinders B cell development (Holmes et al., 2006), demonstrating that PAX5 is not a simple inducer of B cell specific transcript, but it is rather a calibrator which establishes and maintains B cell transcriptional networks (Cobaleda et al., 2007).

Noteworthy, depletion of all transcription factors so far described impairs VDJ recombination, a hallmark of B cell differentiation (Bain et al., 1994; Fuxa et al., 2004; Kee, 2009; Lin and Grosschedl, 1995; Nutt et al., 1997). This process, in fact, consists in laborious rearrangement

of the immunoglobulin gene loci to achieve the production of functional immunoglobulins. Failure in VDJ recombination results in an irreversible developmental block (Schatz and Ji, 2011). RAG enzymes play a central role in this process as they are responsible for recognition, excision and repair of the recombining immunoglobulin regions (Schlissel, 2003). Indeed, *Rag ko* mice showed complete abortion of VDJ recombination (Mombaerts et al., 1992; Shinkai et al., 1992). Altogether, these evidences outline the presence of two important check points during B cell commitment: PAX5 mediated gene regulation and VDJ recombination.

During the pro B cell stage, early VDJ combinatorial events generate a productive heavy chain which is temporary expressed on the cell surface together with the surrogate light chain (SLC) components  $\lambda 5$  and VpreB, forming the pre B cell receptor (pre-BCR) (Melchers, 2005). This event marks pro B to pre B cell transition.

As soon as it reaches the cell surface, the pre-BCR binds bone marrow stromal cell antigens with high affinity, delivering proliferative signal inside pre B cells through the cytosolic domain (Herzog et al., 2009). At the same time the pre-BCR also triggers a negative feedback which represses the transcription of the surrogate chain  $\lambda 5$ , thus preventing further pre-BCR assembly (Parker et al., 2005). This repression is induced by indirect SLP65 mediated activation of *Irf4* and *Irf8* which induce expression of *Ikaros* and *Aiolos* to downregulate  $\lambda 5$  (Lu et al., 2003; Ma et al., 2008; Ochiai et al., 2013; Thompson et al., 2007). Meanwhile pre B cells start rearranging the immunoglobulin light chains, which will replace the surrogate ones to assemble the definitive B cell receptor, (Herzog et al., 2008; Herzog et al., 2009; Ma et al., 2006); generating differentiated B cells. These cells are now ready to leave the bone marrow reaching the target peripheral organ through the blood stream.

T cell development requires thymic environment to occur. For this reason lymphoid progenitors leave the bone marrow and start seeding the thymus (Miller, 1961; Takahama, 2006). Although is not clear whether there is a specific T cell progenitor population already in the bone marrow, CLPs and ELPs are the lymphoid precursors which show T cell potential *in vivo and in vitro* (Bhandoola et al., 2007; Welner et al., 2008). The branching point when T cell separate from B cells is marked by the expression of NOTCH receptors (Radtke et al., 2004). There are four different NOTCH proteins in mammals, NOTCH 1-4 which can bind five different ligands: DELTA-like 1-3-4 and JAGGED 1 and 2 (Hayday and Pennington, 2007). Notch ligands are mainly expressed by thymic epithelial cells (TECs), restricting T cell lineage specification to the thymus (Jenkinson et al., 2006). Interestingly, depletion of *Notch1* induces B cell differentiation in the thymus (Radtke et al., 1999), while its overexpression results in ectopic T cell development in the bone marrow (Pui et al., 1999). These observations indicate that potentially T and B cell can differentiate in either places and that the B versus T

cell choice determination are mainly dictated by the microenvironment cells are sitting in. Because a detailed explanation of T cell developmental mechanisms goes beyond the explanatory scope of this introduction, they won't be further described. However; many aspects of this topic have been reported in the following reviews (Bhandoola et al., 2007; Carpenter and Bosselut, 2010; Hayday and Pennington, 2007; Rothenberg, 2007; Schlenner and Rodewald, 2010; Takahama, 2006).

Although LMPPs still retain myeloid potential (Mansson et al., 2007), common myeloid progenitors (CMPs) are the main source of myeloid and erythroid cells. Definition of CMPs from MMPs marks the other big branching point during hematopoietic development (Akashi et al., 2000). CMPs have the potential to differentiate in granulocytes macrophage progenitors (GMPs) and megakaryocyte erythrocyte progenitors (MEPs), which will terminally differentiate in granulocyte/macrophages and megakaryocytes/erythrocytes, respectively (Akashi et al., 2000; Iwasaki and Akashi, 2007). We have already mentioned that PU.1 dosage in MMPs might be determinant to resolve the lymphoid-myeloid dichotomy (DeKoter and Singh, 2000). Although this concept struggled to be accepted, recent work confirmed that PU.1 mediated myeloid restriction is dosage dependent (Kueh et al., 2013). However; differently from B cells that decrease *Pu.1* expression at the transcriptional level, *Pu.1* transcripts raise in CMPs via lengthening cell cycle duration without direct regulation of gene expression (Kueh et al., 2013). Although high dosage of *Pu.1* marks the branch out from the lymphoid lineage, it has been observed that *Pu.1* deficient mice lack lymphoid and myeloid cells but still develop megakaryocytes and erythrocytes (Dakic et al., 2005; Iwasaki et al., 2005). This implies that PU.1 initiates the myeloid commitment but it is not involved in the specification of the erythroid lineage. Moreover, it has been noted that within CMPs (which have both myeloid and erythroid potential) *Pu.1* expression was heterogeneous. Precisely; CMPs which showed high *Pu.1* levels exhibited myeloid potential, while *Pu.1*-low-expressing cells differentiated in megakaryocytes/erythrocytes. These observations reinforced the idea that *Pu.1* might be important for myeloid but not erythroid lineage specification (Back et al., 2005; Nutt et al., 2005). Consequently, we deduce that either some CMPs immediately downregulate *Pu.1* transcripts to initiate the erythroid program as soon as they diverge from lymphopoiesis, or the erythroid fate is held by other progenitors (Adolfsson et al., 2001; Rosenbauer and Tenen, 2007).

A model system to study myeloid versus MegE choice was provided by the MEPE26 avian cells (Graf et al., 1992; Laiosa et al., 2006). This cell line was established by E26 virus transformation and carries and Myb-Ets oncogene. Upon oncogene activation, these cells differentiated in either erythrocytes or thrombocytes (the avian megakaryocytes), while

induction of RAS or proteinase C pathways resulted in the production of myeloblasts (Graf et al., 1992). Interestingly, the transition to myeloblasts was associated to the downregulation of *GATA1*, a MegE related gene. Enforced expression of *GATA1*, instead, redirected myeloblasts to the erythroid fate and, vice versa, overexpression of the myeloid related gene *PU.1* reprogrammed MegE into myeloblasts (Kulesa et al., 1995; Nerlov and Graf, 1998).

These results indicate that PU.1 and GATA1 antagonize each other's during the myeloid-erythroid diatribe and changes in PU.1:GATA1 ratio is sufficient to redifferentiate cells to the alternative fate (Laiosa et al., 2006). More recent work demonstrated that PU.1 versus GATA1 competition occurs via protein-protein interaction (Nerlov et al., 2000; Rekhtman et al., 1999; Zhang et al., 1999). In fact, PU.1 is able to inhibit GATA1 functions preventing binding to erythroid genes (Stopka et al., 2005; Zhang et al., 2000) and; similarly, GATA1 binds PU.1 preventing transcription of PU.1 related genes (Nerlov et al., 2000; Zhang et al., 1999). GATA1 is; therefore, a crucial transcription factor during erythroid fate determination (Xu and Orkin, 2011). Indeed, *Gata1* deficient mice showed a strong development block during foetal erythropoiesis, which results in embryonic lethality caused by anaemia (Fujiwara et al., 1996; Pevny et al., 1991). These defects can be compensated by *Gata2* and *Gata3* overexpression, probably due to their compensatory role in the absence of *Gata1* (Takahashi et al., 2000). Additionally, impaired expression of *Gata1* specifically in megakaryocytes strongly reduced platelets production and resulted in abnormal megakaryocytic growth (Shivdasani et al., 1997).

Another transcription factor essential for MegE differentiation is FOG1 (friend of GATA1). As the name already suggests; FOG1 teamworks with GATA1 to induce transcription of target genes. FOG1 deficient mice show a phenotype which closely resembles the one observed upon loss of *Gata1* (Tsang et al., 1998; Tsang et al., 1997); however loss of *Fog1* resulted in a stronger phenotype which completely abolished megakaryocytes differentiation (Tsang et al., 1998). As proof of the fact that FOG1 function during megakaryopoiesis is strictly linked to GATA binding, transgenic mice carrying GATA proteins incapable to bind FOG1 have been produced. These animals also lack megakaryocytes, similarly to *Fog1* deficient mice (Chang et al., 2002).

Once MegE precursors are committed, they still have to branch out from each other to terminally differentiate in erythrocyte and megakaryocyte. The candidate proteins which seem to play a role in this regards are the cross-antagonizing factors FLI-1 and EKLF (Starck et al., 2003).

Augmented PU.1 expression together with its interaction with other factors induces myeloid fate, converting CMPs in granulocyte macrophage progenitors (GMPs) (Mak et al.,

2011). This transition is defined by C/EBP $\alpha$ , the transcription factor responsible for GM (granulocytes/macrophages) related-gene activation in combination with PU.1 (Friedman, 2002). Complete C/EBP $\alpha$  depletion leads to postnatal lethality, as mutant mice survived only 8 hours after birth (Wang et al., 1995). Analysis of the hematopoietic compartment in the newborns and in the embryonic foetal liver displayed specific loss of granulocytes, particularly neutrophils and eosinophils (Zhang et al., 1997). This phenotype was confirmed in adult mice where induced loss of C/EBP $\alpha$  abrogated GMP production but did not affect later stages of myeloid differentiation (Zhang et al., 2004), indicating that this factor is not required during late myelopoiesis where, maybe, other C/EBP family members play a role (Laiosa et al., 2006; Zhang et al., 2004). Once CMP to GMP transition has happened; some cells will become granulocytes (eosinophils, basophils and neutrophils) others macrophages. Granulocytes versus macrophage choice seems to be dictated by the transcription factor GFI1, which function is essential for neutrophil differentiation and for shutting down macrophage/monocytes related transcripts (Hock et al., 2003). The remaining granulocytes, basophils (mast cells) and eosinophils, will end their differentiation program by respectively increasing *Gata1* (Nei et al., 2013) and *Gata2* expression (Harigae, 2006; Iwasaki and Akashi, 2007; Orkin and Zon, 2008).

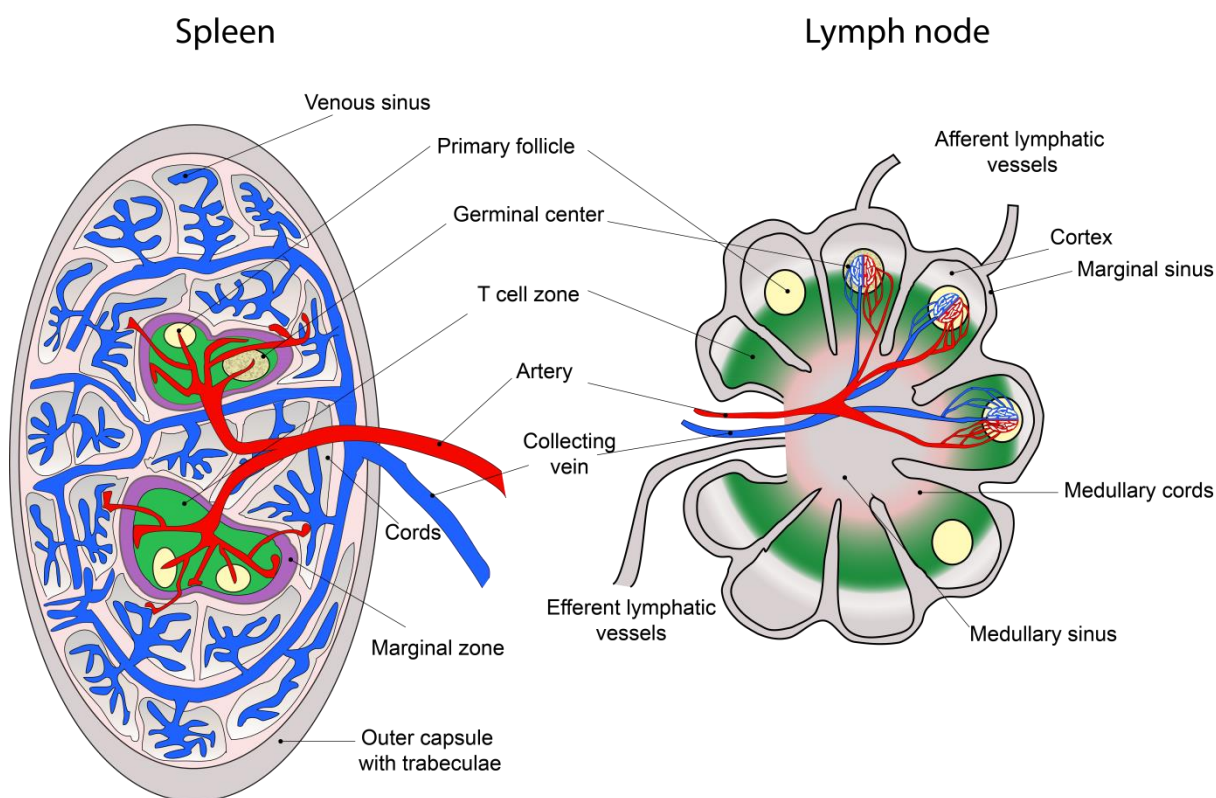
### **1.2.2 Peripheral haematopoiesis**

Mature hematopoietic cells are released into the blood stream from bone marrow and thymus (Janeway, 2008b). Some of them will constantly recirculate in the extracellular fluids as sentinels responsible for body first line defence (innate immunity); others instead will populate peripheral organs to build up the adaptive immune answer. Innate immune answer is mediated by immune effectors like granulocytes, mast cells, macrophages, dendritic cells (DCs), natural killer (NK) cells and soluble factors, such as complement proteins. All together these components have the capability to recognize generic pathogen structures, mediating their demolition and extinction. By contrast, the adaptive immune system enrolls certain cell types, such as B, T and dendritic cells, which are able to build up a specific immune answer to fight those pathogens which bypass the innate recognition barrier (Dranoff, 2004; Janeway, 2008b; Janeway, 2001).

Adaptive immune answer is taking place in peripheral organs, particularly spleen and lymph nodes. These organs are populated by different cells types that remain in a quiescent state until an immune challenge awakens them from dormancy. Peripheral organs architecture is depicted in **Fig. 1.5**. Grossly, it consists of compartmentalized areas called red and white pulp, where

## Introduction

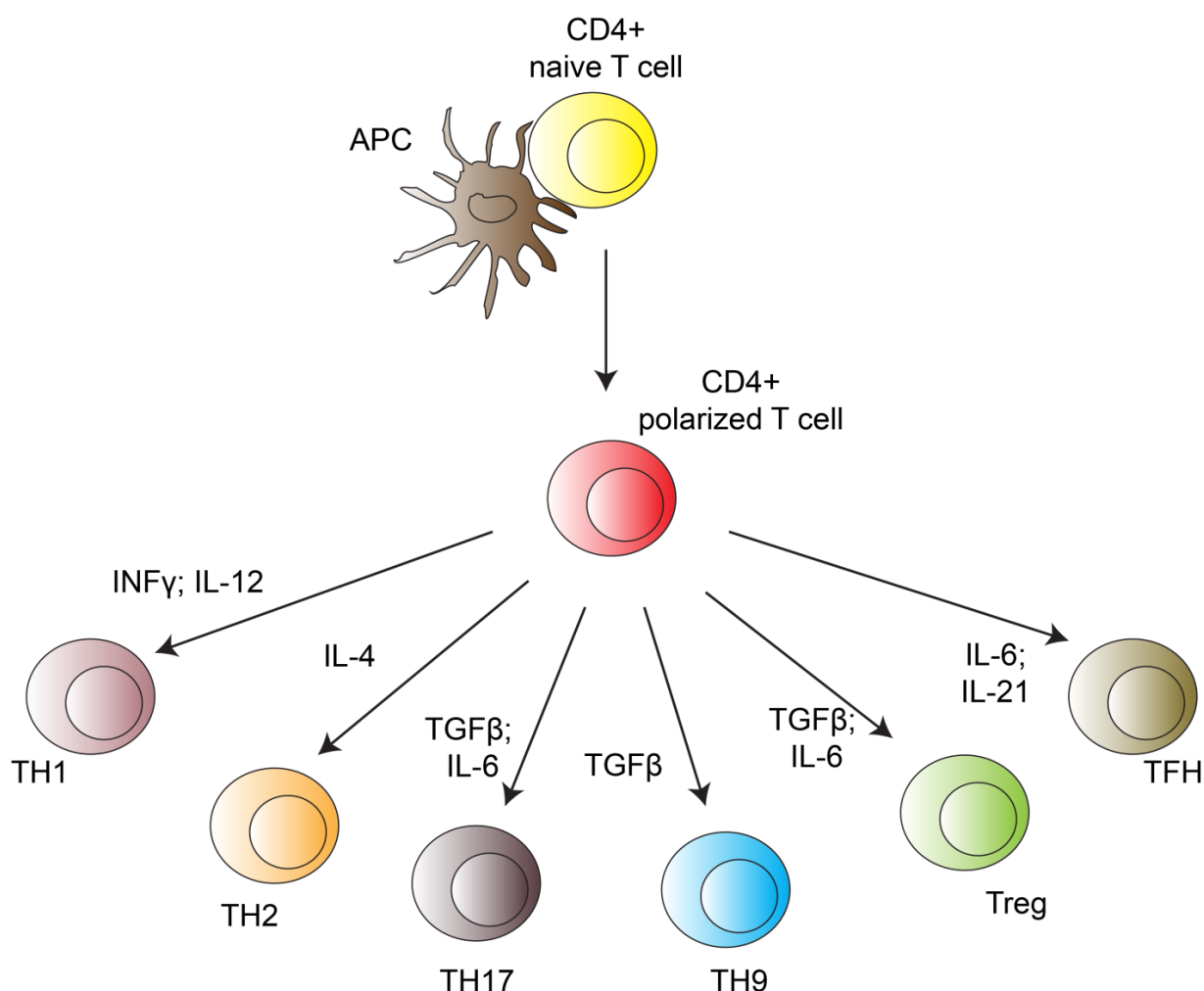
cells are retained by *in loco* produced cytokines. The white pulp is the compartment which includes a layer of T cells, which sits in the T cell zone; and an inner roundish structure called follicle, containing follicular B cells. The red pulp instead, contains erythrocytes and other blood cells (like macrophages, and granulocytes and dendritic cells). In the spleen white pulp and red pulp are separated by the marginal zone, a specialized area which contains B cells and macrophages involved in the innate immunity (Mebius and Kraal, 2005; Mueller and Germain, 2009; von Andrian and Mempel, 2003).



**Figure 1.5 Schematic structures of spleen and lymph node**

Upon immune challenge T and B cells start to proliferate and differentiate in very specialized effectors which are suited to initiate a pathogen specific immune reaction (Janeway, 2008b; Janeway, 2001). For example, T cell encounter with pathogens will induce (1) their activation and (2) their cytokine-instructed differentiation in T helper 1 (TH1), T helper 2 (TH2), T helper 17 (TH17), follicular helper T cells (TFHs) and regulatory T cells (T regs) (Sethi et al., 2013; Swain et al., 2012; Zhou et al., 2009). Remarkably, the instructive role of the cytokines in this context will result in the expression of specific transcription factors, which will polarize CD4<sup>+</sup> dormant T cells to a unique destiny (Lu et al., 2011; Zhou et al., 2009). The

cytokine responsive transcription factors which determine T cells fates are summarized in **Fig. 1.6**.



**Figure 1.6 T helper cell subtypes**

Upon antigen presentation by APCs, naive CD4<sup>+</sup> T cells are polarized to specific fates through the expression of specific transcription factors.

Within the different T helper cells subtypes, TFHs drew the attention of many investigators during the past ten years. Maybe this is due to the special role they fulfil in initiating and maintaining the cell-mediated immune response. The discovery of TFHs dates back to 2005 when Vinuesa and colleagues showed that inactivating mutation in the *roquin* gene, known as sanroque mutation, resulted in the brake out of inflammatory T cells and spontaneous germinal centers formation. This phenotype was caused by the loss of *roquin* repressing function over the costimulatory receptor ICOS, which was thereafter designated as a TFHs hallmark (Vinuesa et al., 2005). The identification of TFHs as a clearly distinguished population which mediates the adaptive immune answer opened a new chapter in the T cell biology field. In fact, many efforts have been made to reveal the molecular mechanisms which sit behind TFHs differentiation and, even if many details are still missing, a lot of results which clarified many



aspects of TFHs commitment and function have been already achieved (Liu et al., 2013; Ma et al., 2012b; Nutt and Tarlinton, 2011; Vinuesa and Cyster, 2011).

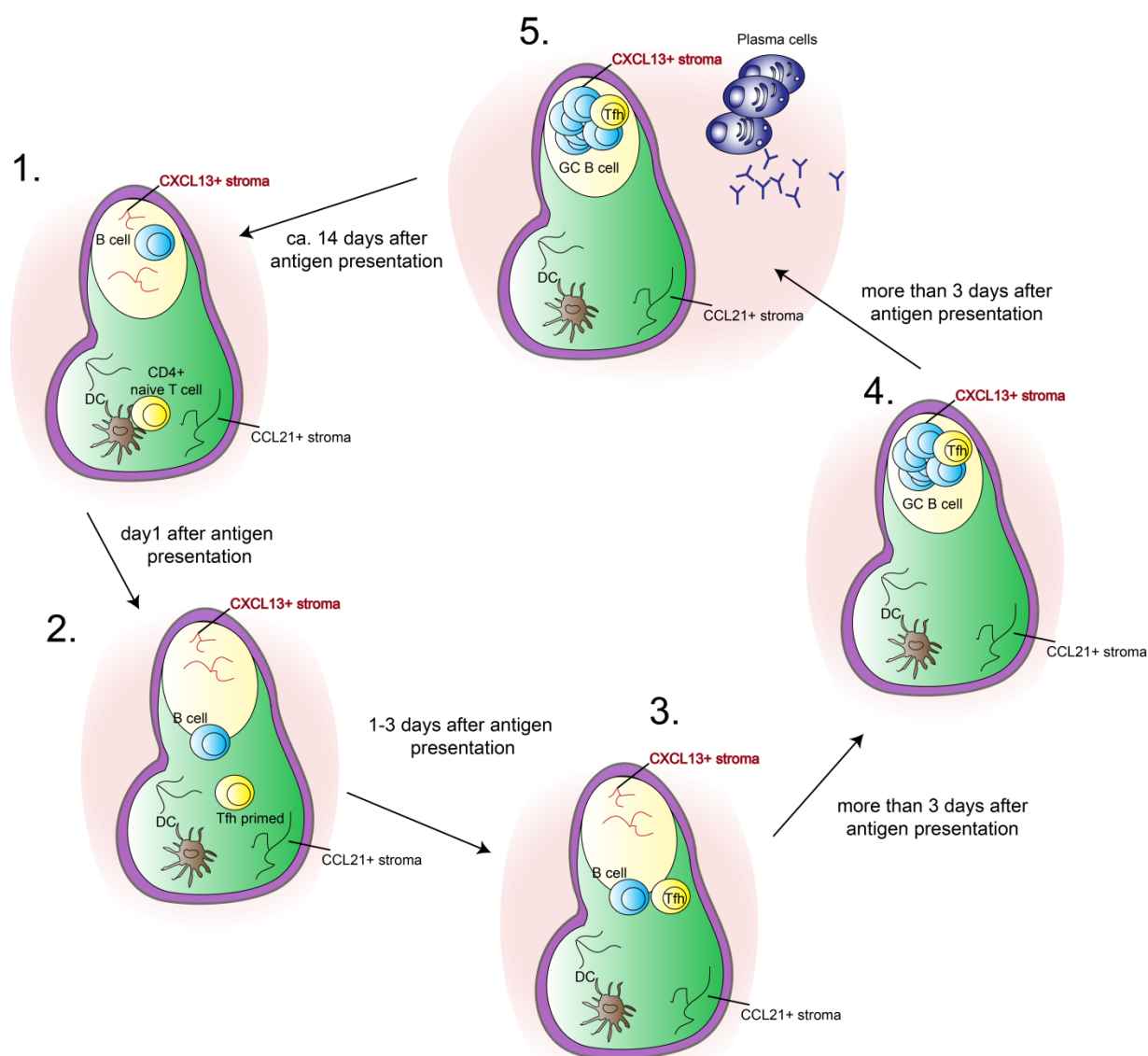
### 1.2.2.1 The germinal center reaction: molecular mechanisms

Adaptive immunity is the way the immune system learns fighting unknown microorganisms and keeps memory of them. In fact, when immune germline receptors fail to recognize incoming pathogens, the body starts to generate a broad spectrum of new receptors seeking the best one out (Janeway, 2008b). This process takes place in secondary lymphoid organs and it is known as germinal center reaction (Nutt and Tarlinton, 2011; Vinuesa and Cyster, 2011).

The priming force which prompts germinal center response is given by antigen presenting cells (APCs). The most specialized APCs are DCs which have the main task to upload on their surface parts of the processed pathogens and present them via cognate interaction to CD4<sup>+</sup> T cells (Vinuesa et al., 2010) (**Fig. 1.7; 1.**).

A special subclass of DCs, called follicular dendritic cells, is able to move around the peripheral organs until reaching the T cell zone (Heesters et al., 2014). Here, T cells will start to be engaged by DC receptors, stabilizing their contact (Celli et al., 2007). At this point T cells are no longer dormant and they start proliferating and transcribing specific genes which will slowly provide them typical TFHs features (Liu et al., 2013).

A very recent discovery identified the imprinted gene Achaete-scute complex homolog 2 (*Ascl2*) as one of the earliest initiator of TFHs commitment. In fact, *Ascl2* deficient mice fail in producing TFHs, preventing germinal center reaction. Molecularly, this was due to the loss of ASCL2-mediated activation of the *Cxcr5* gene (Liu et al., 2014). This gene encodes a surface molecule which is not expressed on naive T cells and is responsive to CXCL13, a cytokine which is produced in the follicular zone. CXCR5R expression on T cell surface starts to mobilize them towards the B cell follicles, their final destination (Hardtke et al., 2005). In homeostatic condition, *Cxcr5* is kept silent because naive CD4<sup>+</sup> T cells express the 17-92 mirna clusters, which targets *Cxcr5* transcripts. CXCR5 expression on TFHs surface, in fact, requires BCL6, the transcriptional repressor that mediates transcriptional silencing of mir17-92 cluster, allowing *Cxcr5* expression (Yu et al., 2009). BCL6 is considered one of the master regulators of TFH cells differentiation (Nurieva et al., 2009). Indeed, its repressive activity inhibits the establishment of TFH cell fate antagonists, as for example the germinal center inhibitor BLIMP1 (Johnston et al., 2009; Shaffer et al., 2000; Tunyaplin et al., 2004).



**Figure 1.7 Schematic representation of germinal center reaction**

1) APCs present the antigen to CD4+ T helper cells committing them to the TFH fate. 2) TFHs leave the T cell zone and start to migrate towards the follicles. 3) Before entering the follicular zone TFHs engage B cells which are located at the periphery of the follicle or B cells which started to migrate outside the follicular zone. This event results in the production of generic immunoglobulin belonging to the IgM and IgG1 class. 4) TFHs enter the follicular zone engaging a strong binding with follicular B cells. This interaction will induce class switch recombination and somatic hypermutation to produce the most specific immunoglobulin against the pathogen. 5) Once the reaction is completed B cells start to differentiate in antibody secreting plasma cells and memory B cells.

Moreover, BCL6 seems to exert inhibitory function by binding promoter regions of *Tbx21*, *Rorc* and *Gata3* genes, which respectively associate with TH17, TH1 and TH2 cell fate (Harris et al., 1999; Nurieva et al., 2009; Yu et al., 2009). The importance of BCL6 during TFH cells differentiation has been broadly demonstrated *in vivo* by using animal models which lacked or overexpressed the transcriptional repressor. As expected, *Bcl6* deficient mice strongly inhibits TFHs differentiation and germinal center formation, while enforced

## Introduction

---

expression of *Bcl6* in CD4<sup>+</sup> naive T cells promoted the establishment of TFHs signature (Fukuda et al., 1997; Yu et al., 2009).

The initial T:DC cells interaction also leads to production of certain cytokines which support TFHs differentiation and maintenance, such as IL-21 and IL-6 (Crotty, 2011; Eto et al., 2011; Nurieva et al., 2008). Their initial activation seems to be controlled by STAT3, a transcriptional regulator in absence of which TFHs are severely reduced, probably due to the aberrant production of TFH related cytokines (Ma et al., 2012a). On the way to the follicles, TFHs will carefully establish and sustain their transcriptional program maintaining high level of IL-21, which transcription is supported by other transcriptional regulator, such as BATF, c-MAF and IRF4 (Bauquet et al., 2009; Betz et al., 2010; Bollig et al., 2012; Ellyard and Vinuesa, 2011; Kwon et al., 2009; Li et al., 2012). While moving towards the follicles (**Fig. 1.7; 2.**), T cells will already encounter antigen-specific B cells which got activated and started to migrate towards the T cell zone due to upregulation of CCR7, a receptor attracted by the T cell zone chemokine CCL12 (**Fig. 1.7; 2. and 3.**). Already at this site, TFHs will start engaging B cells, initiating the extrafollicular reaction, which is also supported by IL-21 (MacLennan et al., 2003; Odegard et al., 2008; Vinuesa et al., 2010). At the next stage, extrafollicular B cells get activated; however they will not be able to produce pathogen specific immunoglobulin, but only low-affinity antibodies with low level of hypermutation (MacLennan et al., 2003). Now, TFHs are about to enter the follicular area as CXCR5 high, ICOS high, PD-1<sup>+</sup> cells which can be clearly discriminated by FACS from the other CD4<sup>+</sup> T cells. Terminally differentiated TFHs are now ready to engage follicular B cells (Vinuesa and Cyster, 2011) (**Fig. 1.7; 4.**). This interaction will last a sufficient amount of time to induce the rearrangement of new, more specific immunoglobulin via two processes known as class switch recombination and somatic immunoglobulin hypermutation (Kinoshita and Honjo, 2001). Importantly, the strength of the B:T cells binding will represent an initial parameter to exclude those B cells which are weakly compatible to T cell surface marker. A strong binding will provide a survival signal which will allow production of new specificity while, at the opposite, a feeble interaction will induce cell death to avoid likely generation of unwanted self-reactive antibodies (Klein and Dalla-Favera, 2008; Nutt and Tarlinton, 2011).

Activated germinal center B cells also require the establishment of specific transcriptional networks which prevent their premature differentiation in antibody secreting plasma cells. In this regard, several transcriptional regulators have been identified as crucial for the germinal center B cells maintenance. They include PAX5, MITF, BCL6, MTA3 and BACH2 (Shapiro-Shelef and Calame, 2005).

Beyond its function during early B cell development, PAX5 is also essential for later stage of B cell development (Horcher et al., 2001; Nutt et al., 2001). During germinal center formation, this master regulator exerts activating function on the activation-induced cytidine deaminase (AID), the enzyme which generates a new antibody repertoire mediating somatic hypermutations and class switch recombination processes (Gonda et al., 2003). Additionally, PAX5 also represses the X-box-binding protein 1 (*Xbp-1*), an essential factor which activates the plasma cells secretory program (Reimold et al., 1996). Another positive regulator of germinal centers development is microphthalmia-associated transcription factor (MTIF). In the absence of MTIF, B cells induce *Blimp1* expression and undergo extemporaneous production of plasma cells (Lin et al., 2004). Probably, this is due to *Irf4* upregulation which promotes plasma cell differentiation. Notably, IRF4 has been found to be important either for germinal center formation or for plasma cell specification. This double function connects with dosage-dependent activation of the *Irf4* gene, as low and high *Irf4* expression levels respectively associate with germinal center B cell and plasma cell differentiation (Ochiai et al., 2013).

The exhaustion of the germinal center reaction coincides with an increase of *Prmd1*, the gene which encodes BLIMP1 (Kallies et al., 2004). This protein is an inhibitor of germinal center reaction and is required for plasma cell differentiation (Shapiro-Shelef et al., 2003). In fact, BLIMP1 is a zinc finger protein which activates plasma cells related genes (Sciammas and Davis, 2004; Shaffer et al., 2004) as *Blimp1* overexpression has been shown to induce spontaneous plasma cell differentiation (Turner et al., 1994). Direct *Blimp1* repression is mediated by BCL6 which is highly expressed in germinal center B cells (Cattoretti et al., 1995; Tunyaplin et al., 2004). Indeed, BCL6 repressive function is essential for germinal center formation since *Bcl6* deficient B cells do not form any germinal center (Dent et al., 1997). Specifically, BCL6 mediates direct repression of *Blimp1* via antagonizing the activity of AP-1 and STAT3 transcriptional activators (Reljic et al., 2000; Vasanwala et al., 2002). Moreover, BCL6 repressive function on *Blimp1* promoter was also found to be mediated by MTA3 (metastasis-associated 1) interaction, as MTA3 knock-down does not affect *Bcl6* expression but activate plasma cell specific genes (Fujita et al., 2004). Besides MTA3; BACH2 has also been found to interact with BCL6. Genome-wide analysis showed that they cooccupy the cis-regulatory sequence of *Blimp1*, likely inactivating it (Huang et al., 2014). *Bach2* deficient mice also fail to form germinal centers and they do not activate AID (Muto et al., 2004).

Extinction of the germinal center response will induce plasma cell differentiation through activation of appropriate genes, as *Blimp1*, *Xbp-1* and *Irf4* (Shapiro-Shelef and Calame, 2005). Newly formed plasma cells will either survive as antibody-secreting cells or will persist as

memory B cells, the archive of the immune system (McHeyzer-Williams et al., 2012; Radbruch et al., 2006)(Fig. 1.7 5.).

### 1.3 The histone methyltransferase SETDB1

SETDB1 and SUV39H are the histone methyltransferases which establish H3K9 tri-methyl marks (O'Carroll et al., 2000; Rea et al., 2000; Schultz et al., 2002). Although they associate with the same histone modification, these enzymes work in different part of the genome. SUV39H-dependent H3K9me3 is highly enriched at pericentromeric region, while, outside these areas, the same methyl mark seems to be established by SETDB1 (Bilodeau et al., 2009; Dambacher et al., 2010; Karimi et al., 2011). The methyltransferase activity of SETDB1 is given by the presence of a pre-SET, SET post-SET domain; while the Tudor domain and MBD (methyl-CpG-binding) domain mediate protein-protein interaction and binding to methylated DNA, respectively (Chen et al., 2011; Kang, 2014) (Fig. 1.8).

#### SETDB1



**Figure 1.8 The histone methyltransferase SETDB1**

Remarkably, *Setdb1 ko* mice die at perimplantation (Dodge et al., 2004), demonstrating that SETDB1 mediated silencing is essential at very early stages of mouse development. Unfortunately, the fact that *Setdb1 ko* embryos die before implantation hinders the possibility to perform *in vivo* analysis. To circumvent this issue, mouse models which conditionally delete the protein in adult tissue have been generated. Consistently with the phenotype observed in *Setdb1 ko* embryo, deletion of the enzyme during adult development strongly interferes with cell differentiation. For example, conditional ablation of *Setdb1* alleles in the brain leads to decreased H3K9me3 and aberrant gene expression, resulting in severe brain defects and early lethality (Tan et al., 2012). More recently, a developmental role of SETDB1 has been identified in the context of bone development. Here, mesenchyme specific deletion of *Setdb1* causes ectopic hypertrophy, apoptosis and terminal differentiation of articular chondrocytes, predisposing articular cartilage to degeneration (Lawson et al., 2013b). Intriguingly, using the same knockout scheme, the same group also showed that lack of

SETDB1 inhibits differentiation of mesenchymal stem cells in osteoblasts (Lawson et al., 2013a).

Because of technical limitations and/or lack of material, most of the *in vivo* SETDB1-related phenotypes are carefully described but poorly explained at the molecular level. Thus, up to now, all experiments to elucidate SETDB1 molecular functions have been performed *in vitro*, mostly in mESCs.

In 2002, Schultz *et al.* revealed that in NIH/3T3 cells SETDB1 mostly localizes at euchromatic regions, where it binds the corepressor KAP1 (KRAB associated protein 1) (Schultz et al., 2002). KAP1 recruitment at euchromatin is mediated by KRAB-ZNF proteins, which specifically recognize their corresponding cis-regulatory elements across the DNA (Iyengar et al., 2011). KAP1 binding to KRAB-ZNFs is an obligated step to induce transcriptional repression (Gröner et al., 2010; Sripathy et al., 2006). In fact, this protein will become the anchor bolt for the assembly of several repressive complexes which will ultimately mediate gene silencing (Friedman et al., 1996; Schultz et al., 2002). The identification of SETDB1 as a novel transcriptional repressor that localizes at euchromatin to induce gene silencing, arouse a great deal of interest in understanding which regions in the genome are SETDB1 targets. The first study in this direction was performed in mouse embryonic stem cells (mESCs). Here, high-throughput shRNA screening approach spotted SETDB1 as one of the most important chromatin regulator which stabilizes mES cell state. Interestingly, following genome-wide analysis showed that SETDB1 and H3K9me3 cooccupied developmental genes, possibly inhibiting their transcription (Bilodeau et al., 2009). These findings launched SETDB1 as the candidate chromatin modifier, responsible for silencing lineage inappropriate genes during development transitions. However, more recent publications challenged this belief; providing convincing experimental evidences that this histone methyltransferase mainly represses endogenous retroviral elements (ERVs), a class of proviral sequences stably integrated in the genome (Karimi et al., 2011). Indeed, it has been shown that *Setdb1* deficient mES cells highly transcribed ERVs (in particular IAP elements), and that these elements are directly repressed by SETDB1. Interestingly, contemporary loss of DNMTs, minimally affected the transcriptional activity over these regions, indicating that the pathway which mediates their silencing is DNA methylation-independent (Karimi et al., 2011). Because SETDB1 interacts with the co-repressor KAP1, ChIP-Seq experiments have been performed to investigate whether KAP1 also binds ERVs genomic regions, mediating their silencing. Notably, in ES cells, KAP1 localized on ERVs LTR promoters, supporting the idea that these regions are mainly silenced by the KRAB-ZNFs/KAP1 repression system. Consistently with these findings, KAP1 deficient mES cells significantly upregulated ERVs (especially IAP elements),

demonstrating that KAP1 is also necessary for ERVs silencing (Rowe et al., 2010). Although the presence of SETDB1 and KAP1 seems to be a constant for retrotransposons silencing, the KRAB-ZFN proteins which mediate KAP1 recruitment might be specific for specific classes of retrotransposons. ZFP809, for example, has been identified as one of the KRAB-ZFN which specifically binds retroviral DNA sequences inducing *de novo* silencing upon KAP1 recruitment. This occurs via ZFP809 recognition of the primer binding site (PBSs) complementary to proline tRNA (PBSPro), which is specifically contained in a class of elements called murine leukaemia virus (MLVs) (Wolf and Goff, 2009). However, other endogenous retroviruses contain distinct PBSs which are recognized by other ZFPs, indicating that KAP1 recruitment is mediated by alternative KRAB-ZNFs which are retrotransposon specific (Rowe and Trono, 2011).

Although we are gaining more and more knowledge about SETDB1 functions; it remains a contentious whether it is a repressor of developmental gene, a main retrotransposon silencer, or both. Moreover, all *in vivo* phenotypes resulting from *Setdb1* depletion are broadly described but not yet clarified at the molecular level. In conclusion, the reasons why SETDB1 is so essential during development are not yet elucidated and they still represent an open challenge for people working in this field.

#### 1.4 The histone methyltransferases SUV420H

H4K20 tri-methyl marks mainly accumulate at chromatin regions which have to be silenced. In MEFs, SUV420H1 and SUV420H2 have been identified as the histone methyltransferases that mainly act as nucleosomal H4-K20 trimethylating enzymes. In fact, knock-down of both *Suv420h1* and *Suv420h2* selectively impairs H4K20 trimethylation at constitutively silenced chromatin (heterochromatin), where this histone mark is typically acquired together with H3K9me3 (Dambacher et al., 2013; Schotta et al., 2004). The importance of H4K20me3 in heterochromatin formation has been recently underlined in our group. We observed that H4K20me3 is essential for pericentromeric recruitment of choesin, a protein complex which ensures correct separation of sister-chromatids, avoiding mitotic defects due to aberrant chromosome segregation (Hahn et al., 2013).

Beyond playing a structural role in heterochromatin formation, H4K20me3 is also essential during development. *Suv420h ko* mice, in fact, show perinatal lethality, demonstrating that this repressive mark is required during postnatal development. Additionally, conditional deletion of these enzymes in the hematopoietic system impaired pro to pre B cell transition and resulted in defective immunoglobulin class switch recombination (Schotta et al., 2008).

## Introduction

---

More recently, further evidences regarding the importance of H4K20me3 during development have been provided in *Xenopus*. In this regard, Nicetto *et al.* demonstrated that morpholino-mediated disruption of *suv420h1* and *suv420h2*, resulted in neuroectoderm differentiation defects. This failure was consequent to ectopic expression of the pluripotency gene Oct25, a direct target of Suv420h mediated gene repression (Nicetto et al., 2013).

Although SUV420H functions are not extensively elucidated in the context of development, it seems that H4K20me3 is required for proper cell differentiation.



## 2 Results (I)

### 2.1 Spontaneous germinal center activation in *Suv420h2* knockout (ko) mice

#### 2.1.1 *Suv420h2* ko mice show enlarged spleen

To better understand the role that H4K20me3 plays in the context of immune regulation, we used *Suv420h2* ko (d/d) mice as model system and compared them to wild type mice (+/+). To ensure that both genotypes had the same genetic background, we first crossed +/+ mice with d/d mice. This breeding scheme generated 100% of heterozygous mice (+/d), which were then intercrossed, producing in the same littermate +/+, d/d and +/d mice, with frequencies of 1/4, 1/4 and 1/2, respectively (**Fig. 2.1 a**). *Suv420h2* ko mice appeared normal and were born following regular Mendelian ratio.

Because we wanted to investigate the role that the repressive chromatin mark H4K20me3 plays in the immunological context, we started examining the main hematopoietic organs. Although bone marrow cellularity and thymus size were normal, we noticed that *Suv420h2* ko spleens were enlarged. *Suv420h2* ko spleens weighted nearly 30-40% more than wild type spleens and the organ cellularity was also comparably increased (**Fig. 2.1 b and c**). Increased spleen size can be explained by developmental defects; therefore we investigated whether hematopoietic populations were affected in bone marrow and thymus, the central hematopoietic organs where B, myeloid-erythroid and T cells develop, respectively. Immunophenotyping of the main cell populations in these organs showed that all developmental transitions were intact and that each population was generated with normal frequency in the absence of *Suv420h2* (**Table1**).

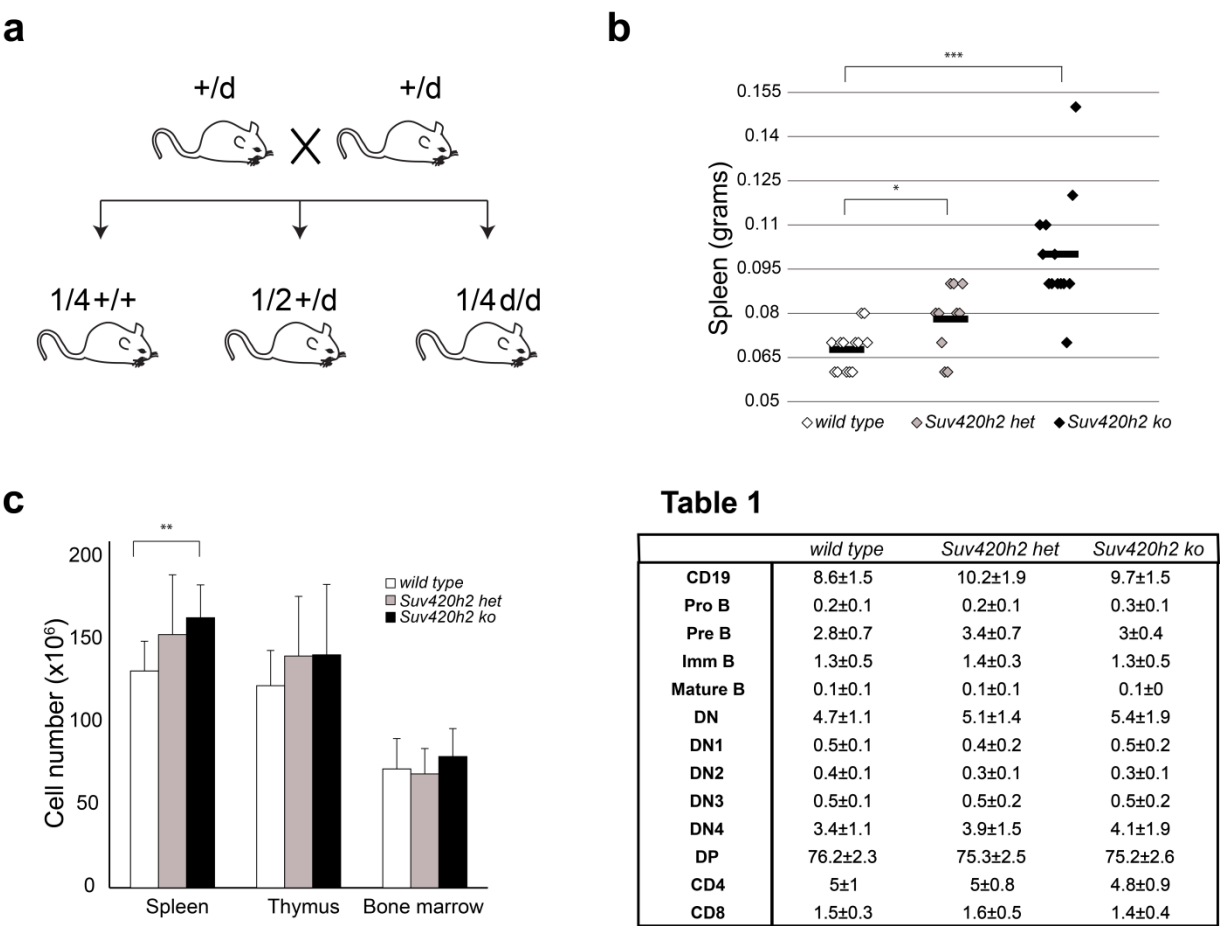
#### 2.1.2 *Suv420h2* ko B cells spontaneously initiated the germinal center reaction

Because any defect during early immune development could explain spleen enlargement, we investigated all splenic populations in a cohort of animals which included 17 to 22 weeks old mice. Spleen is a peripheral hematopoietic organ which mainly contains lymphocytes; therefore we started examining B and T cells using general markers. CD19 was used to detect B cells, while CD4 and CD8 were used to discriminate helper from cytotoxic T cells. Surprisingly, *Suv420h2* T and B cells frequencies were normal when compared to wild type mice (**Fig. 2.2 b**).

Results (I)

Even though the spleen mainly contains B cells, staining with general markers might not be sufficient to discriminate all subpopulations, therefore we started examining more in detail all B cell compartment.

IgD and IgM markers were used to test whether B cell development was affected in the spleen. As shown in **Fig. 2.2 c and d** immature, transitional and mature *Suv420h2 ko* B cells resulted to be unaltered.

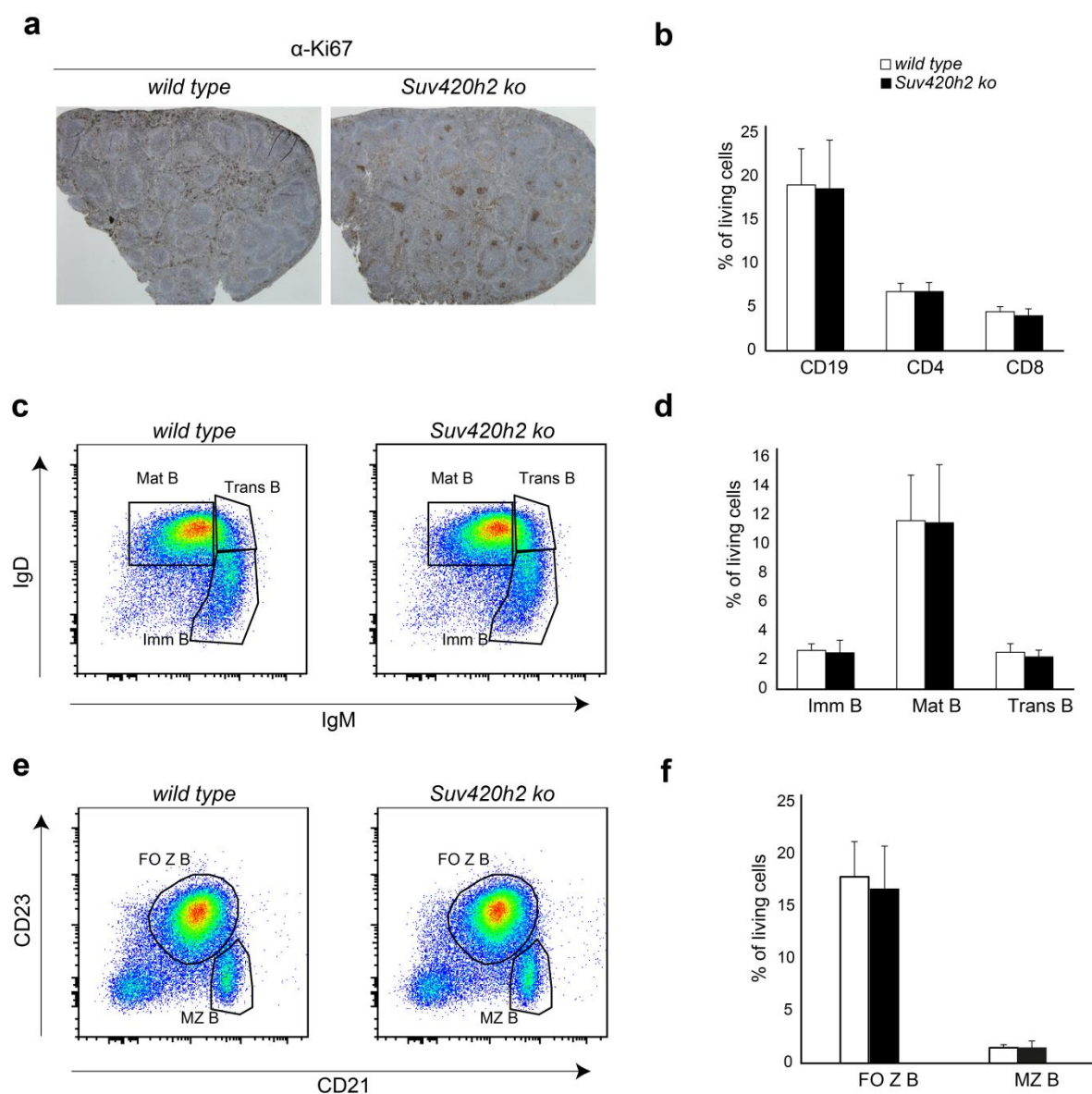


**Figure 2.1 and Table1** *Suv420h2 ko* mice show enlarged spleen and unaltered central haematopoiesis  
**a)** Breeding scheme adopted to obtain wild type and *Suv420h2 ko* mice with the same genetic background. **b)** Spleen weight (n=13) and **c)** hematopoietic organs cellularity from wild type, *Suv420h2 het* and *Suv420h2 ko* mice (n=13). \*\*\*P < 0.001, \*\*P < 0.01, \*P < 0.05 (unpaired two-tailed Student's t-test). **Table1.** B and T cells population frequencies during development. Numbers represent percentage of living cells with related standard deviations.

Within the mature B cell population, follicular and marginal zone B cells can be further distinguished. These, are very specialized cells involved respectively in cell-mediated and innate immunity. Marginal zone (B220+, CD21 high, CD23 low) and follicular B cells (B220+, CD23+, CD21 low) were then tested in *Suv420h2 ko* spleen, but again any difference could be detected in cell numbers or markers distribution (**Fig. 2.2 e and f**).

## Results (I)

Further analysis were also performed to investigate granulocytes, macrophages, dendritic cells, natural killer and erythroid populations but none of them resulted to be impaired (data not shown). All these data are supported by the fact that *Suv420h2 ko* spleen architecture is unaltered (**Fig. 2.2 a**); yet they still do not explain why we observe spleen enlargement in the absence of *Suv420h2*.



**Figure 2.2 Main splenic B cell populations are unaltered in *Suv420h2 ko* mice**

**a**) Microscopic picture showing spleen architecture. Slides were stained with the proliferation marker ki67 **b**) Splenic B (CD19+) and T (CD4+ or CD8+) cell populations (n=6). **c**) Flow cytometry analysis of B cell development in the spleen. Immature (B220+IgM+ IgD-/low), trans B (B220+, IgD+, IgM+) and mature B (B220+, IgD++, IgM+/low) cells (n=6). **d**) Statistics form c). **e**) Marginal zone (B220+, CD21+, CD23-/low) and follicular zone (B220+, CD21-/low, CD23+) B cells (n=6). **f**) Statistics form e).

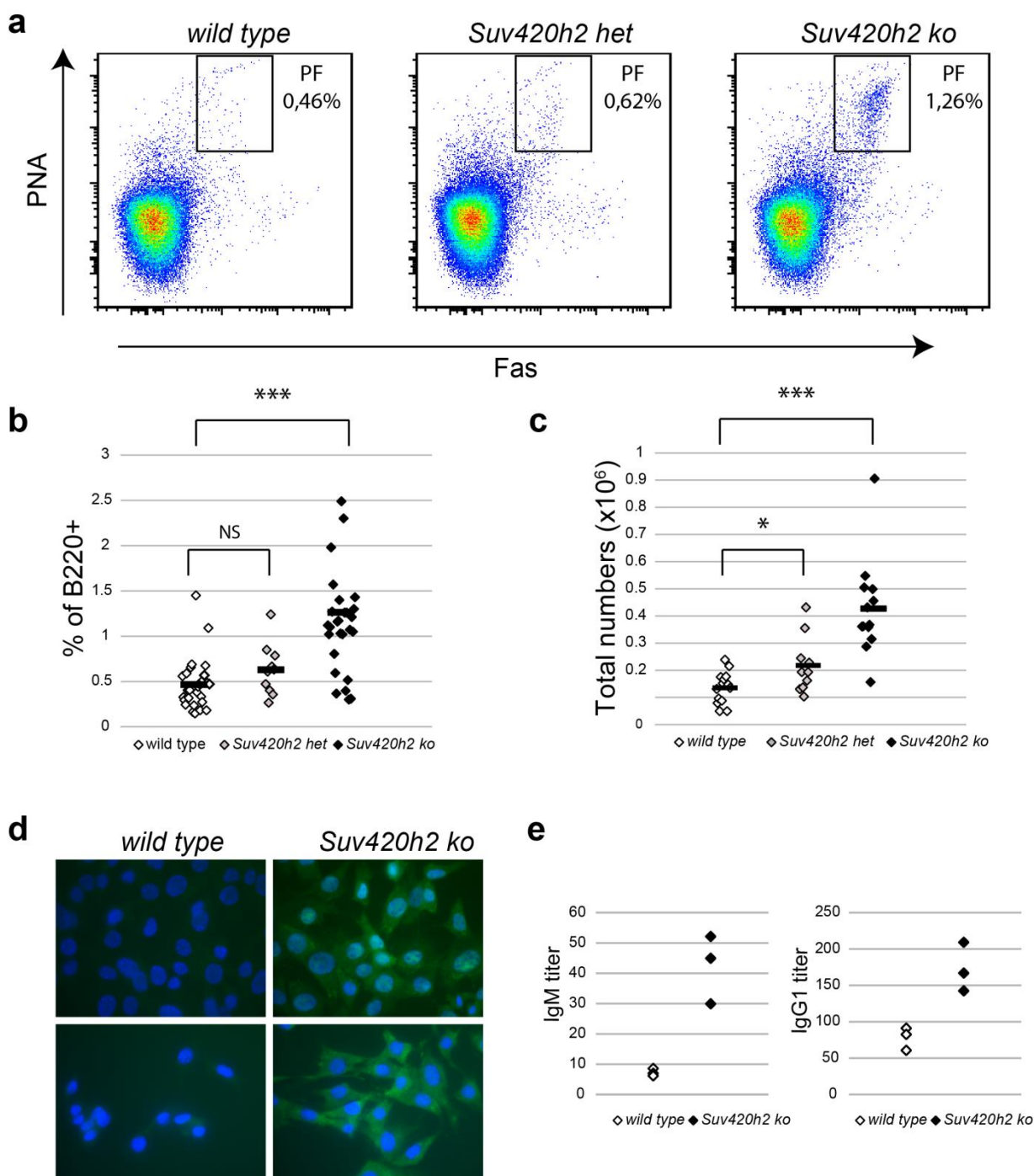
## Results (I)

---

Increased organ size and cellularity correlate with increased cellular proliferation. To understand whether this was the case in *Suv420h2 ko* spleens, we tested spleen sections for presence of the proliferation markers ki67. *Suv420h2 ko* spleens resulted highly positive for ki67, confirming the hypothesis that some of the *Suv420h2 ko* splenocytes are highly proliferative (**Fig. 2.2 a**). Notably, ki67+ cells were mostly located in the follicles, regions of the spleen where follicular B cells sit, although we also observed extrafollicular proliferation. Follicles are spleen areas where, upon immune challenge, germinal center reaction takes place. When germinal center reaction occurs, follicles start to proliferate, generating the very specialized germinal center B cells, which main task is to produce the most specific antibodies against pathogens. These cells have the distinctive feature to express on their surface PNA and FAS (CD95), two typical germinal center markers. When we checked whether *Suv420h2 ko* B cells were positive to these markers, we noticed that a PNA+, FAS+ B cell population popped up in our FACS plot, while in wild type mice the same population was not present (**Fig. 2.3 a**). Remarkably, *Suv420h2 het* mice showed an intermediate germinal center activation; in fact, in average, the percentage of PNA+ FAS+ cells in heterozygous mice was in between wild type and *Suv420h2 ko*, indicating that the observed phenotype correlated with the dosage of *Suv420h2* (**Fig. 2.3 a-c**).

Germinal centers get activated only in response to immune stimuli. These are mainly peptides derived from processing of pathogens, which are first presented by DCs to the TFH cells, which will transfer the information to B cells, inducing their proliferation and priming production of very specific antibodies.

Because *Suv420h2 ko* mice show germinal center activation in the absence of immune stimuli, they are likely to produce autoreactive antibodies which might recognize self-antigens, a phenomenon which often associates with autoimmunity. When antibodies start to be produced, they are released in the blood stream; therefore if *Suv420h2 ko* B cells undergo spontaneous germinal center reactions, the randomly produced antibodies can be detected in the mouse sera. In order to test this hypothesis, we isolated serum from blood of wild type and *Suv420h2 ko* mice and tested it on mouse embryonic fibroblasts (MEFs) for the presence of self-reactive antibodies. Autoreactivity was detected using an anti-mouse secondary antibody conjugated with a green fluorochrome. *Suv420h2 ko* sera recognized and bound many self-antigens expressed by MEFs, as shown by the strong green signal detectable; while wild type antibodies hardly recognized self-antigens (**Fig. 2.3 d**).



**Figure 2.3 Germinal center activation in *Suv420h2* ko spleen**

**a)** FACS plots showing germinal center B cells (B220+, PNA+, FAS+). **b)** Statistics of germinal center B cells calculated as % of the B220+ cells (n=27). NS, not significant and \*\*\*P < 0.001 (unpaired two-tailed Student's t-test). **c)** Total numbers of PNA+ FAS+ cells (n=13). \*\*\*P < 0.001 and \*P < 0.05 (unpaired two-tailed Student's t-test). **d)** MEFs stained with wild type or *Suv420h2* ko sera. Secondary detection was performed using Alex488 antibody. **e)** Preliminary ELISA results to test IgM and IgG1 titres. ELISA was performed in collaboration with Strobl laboratory (Helmholtz Zentrum, Munich).

Mouse models for autoimmunity, in general, produce self-reactive antibodies which bind to very specific antigens and this specificity allows categorizing, in most of the cases, the type of autoimmune disease. Instead, *Suv420h2* ko mice recognized a broad spectrum of antigens,

## Results (I)

---

from the nucleus to the cytosol, suggesting that whatever signal is inducing autoantibodies production, it does not results in generation of immunoglobulins with unique specificity.

Although we proved the existence of autoreactive antibodies in *Suv420h2 ko* mice, we still lack understanding about the amount of immunoglobulins produced and the class they belong to. To better address this point we performed a preliminary ELISA test for the most common class of immunoglobulins (IgM, IgG1, IgG2a, IgG2b). *Suv420h2 ko* mice showed a significant increase in IgM and IgG1 production, whereas IgG2a and IgG2B were not significantly changed (**Fig. 2.3 e and data not shown**). IgM and IgG1 are produced in substantial amounts as soon as the germinal center reaction starts upon B cell encounter with TFHs; while the other two classes of immunoglobulins, IgG2a and IgG2b, are produced upon activation of Th1.

All together these data indicate that loss of *Suv420h2* results in spontaneous activation of germinal center B cells, thus lack of repressive mark H4K20me3 make cells more prone to switch to an active state.

### 2.1.3 *Suv420h2 ko* B cells response to external stimuli is comparable to wild type cells

To verify whether *Suv420h2 ko* B cells can be activated more easily than wild type, we cultured splenic B cells *in vitro* using several kinds of immunostimulants to trigger all different pathways which lead to B cell proliferation.

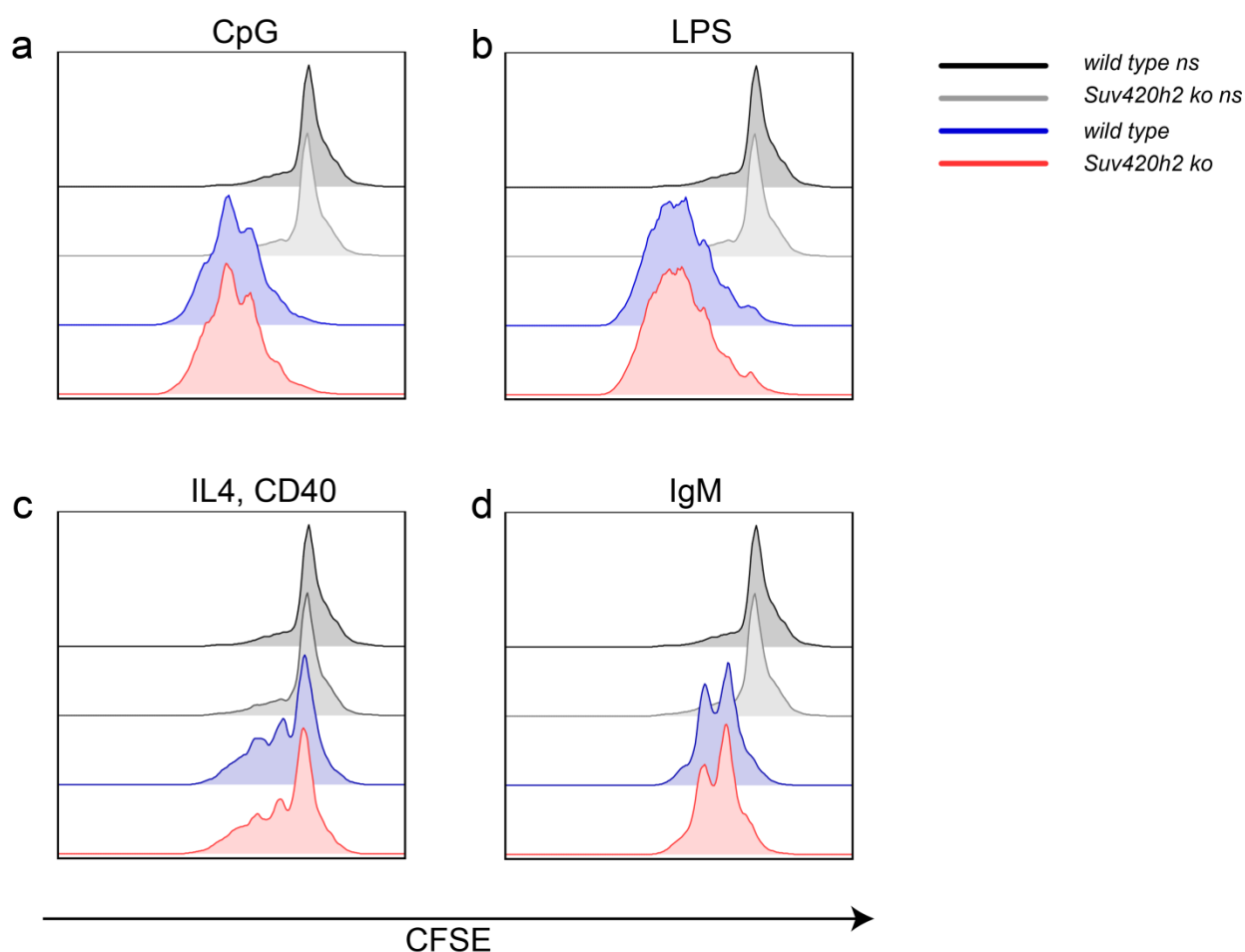
Before stimulation, B cells were stained with carboxyfluorescein (CFSE). This molecule is incorporated by the cells and retained in the cytosol. Cells which go through cell cycle will dilute the dye by half and every division steps will be visualized as a clear peak in a FACS histogram.

After 3 days of stimulation, B cells were analysed by FACS. All dead cells were excluded by staining with the viability dye 7-Aminoactinomycin (7AAD) and the living ones were checked for CFSE dilution. As it shown in (**Fig. 2.4**), B cells responded to all provided stimuli in a way which was comparable to wild type B cells.

Additionally, to investigate changes in gene expression we performed microarray analysis on bulk B cells populations. All deregulated genes were only moderately affected and unfortunately, none of them could be directly linked to the increase in proliferation we observed *in vivo* (**Appendix, Results I**).

Because the analysed population was a mixture of different cell types, which might have different networks of gene regulation, we also looked at candidates which showed minor changes in gene expression. Also in this case, none of them could be attributed to the observed phenotype.





**Figure 2.4** *In vitro* stimulated *Suv420h2 ko* B cells behave comparably to wild type cells.

a) b) c) and d) MACS depleted splenic B cells stained with CFSE were cultured in the presence of different immunostimulants: CpG (10  $\mu$ M); LPS (5  $\mu$ g/ml); IL4 (10 ng/ml) CD40 (2,5  $\mu$ g/ml) and IgM (10  $\mu$ g/ml). After 3 days cell were analysed by FACS to test CFSE dilution. Each peak corresponds to one cell division.

#### 2.1.4 Follicular helper T cells are overrepresented in *Suv420h2 ko* spleen

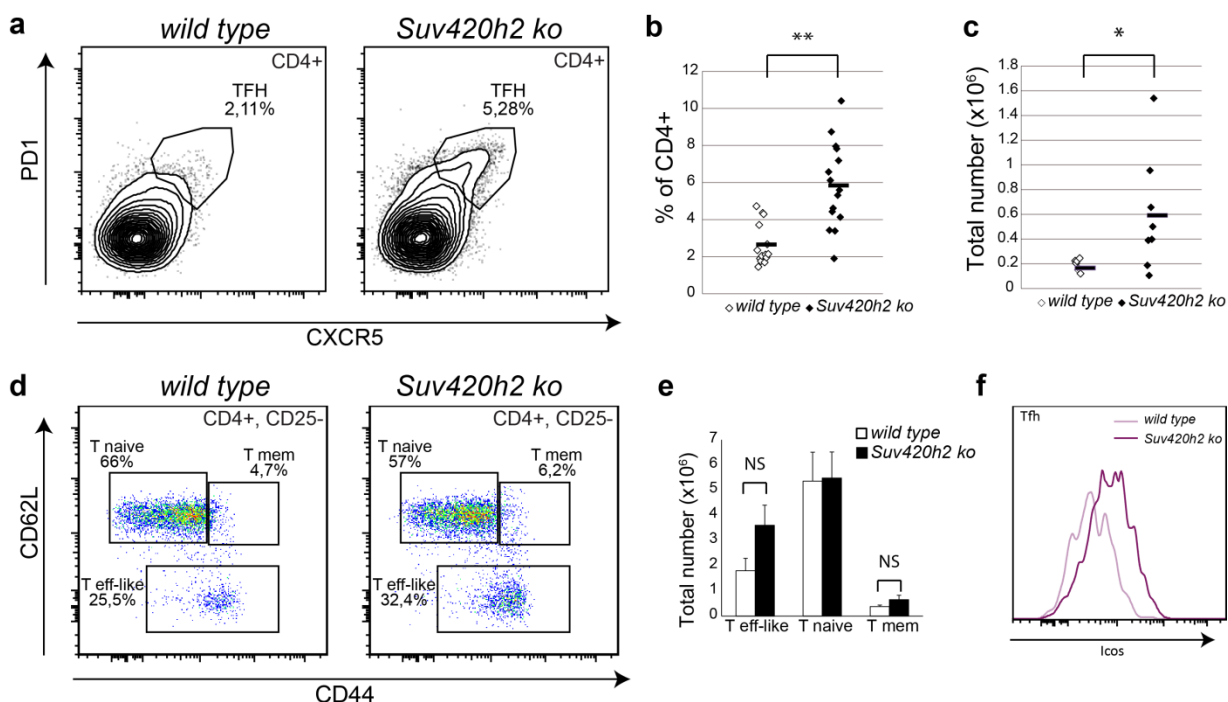
Our *in vitro* data indicated that spontaneous germinal center activation in *Suv420h2 ko* mice seemed not be dependent on intrinsic defects inside B cells; therefore other possibilities had to be explored to understand which mechanisms is corrupted in *Suv420h2 ko* spleens.

IgM and IgG1 antibodies, which are increased in *Suv420h2 ko* mice, are linked to initial stages of germinal center activation. Because this reaction strongly depends on the interaction between TFHs and quiescent B cells, we investigated whether follicular helper T cells were affected. TFHs are newly discovered T cells which have specific gene signatures compared to other T populations and a still not fully understood differentiation pathway. These cells differentiate from CD4<sup>+</sup> naive T cells upon interaction with the dendritic cells which present them the antigens. This cooperation prompts transcriptional changes which allow the expression of a group of genes which promote migration of TFHs from the T cell zone to the follicles. Once TFHs have reached the follicular zone they are fully differentiated and are

## Results (I)

clearly distinguishable from other T cells due to CXCR5, CD4, PD1, and ICOS expression on their surface.

When we tested TFHs in *Suv420h2 ko* mice, we found that these cells were present in the spleen with higher frequencies compared to the wild type. Wild type CD4<sup>+</sup> cells contained around 2% TFH cells. Within *Suv420h2 ko* CD4<sup>+</sup> population, instead, this percentage was increased more than two folds (**Fig. 2.5 a and b**).



**Figure 2.5 TFHs are expanded in *Suv420h2 ko* spleens**

**a)** TFHs (CD4<sup>+</sup>, CXCR5<sup>+</sup>, PD1<sup>+</sup>) population in wild type and *Suv420h2 ko* spleens. **b)** TFHs frequencies shown as percentage of CD4<sup>+</sup> cells (n=13). \*\*P < 0.01 (unpaired two-tailed Student's t-test). **c)** Splenic TFHs absolute numbers (n=7). \*P < 0.05 (unpaired two-tailed Student's t-test). **d)** Resting naive T cells (CD4<sup>+</sup>, CD62L high, CD44<sup>-</sup>/low, CD25<sup>-</sup>) cells, T effector-like (CD4<sup>+</sup>, CD44<sup>+</sup>, CD62L low, CD25<sup>-</sup>) and memory T cells (CD4<sup>+</sup>, CD44<sup>+</sup>, CD62L<sup>+</sup>, CD25<sup>-</sup>) frequencies. **e)** Total cell number of all T cell subpopulation shown in d) (n=4) NS, not significant (unpaired two-tailed Student's t-test). **f)** FACS histogram depicting ICOS expression on TFHs surface.

Importantly, all *Suv420h2 ko* TFHs expressed high level of ICOS, the main follicular T cells marker, confirming cell identity (**Fig. 2.5 f**). Additionally; we also checked whether other T cells populations were affected by increased presence of TFHs in *Suv420h2 ko* spleens. We stained CD4<sup>+</sup> cells with CD25, CD44 and CD62L. These markers are used to differentiate resting naive T cells (CD4<sup>+</sup>, CD62L high, CD44<sup>-</sup>/low, CD25<sup>-</sup>) cells, T effector-like (CD4<sup>+</sup>, CD44<sup>+</sup>, CD62L low, CD25<sup>-</sup>) and memory T cells (CD4<sup>+</sup>, CD44<sup>+</sup>, CD62L<sup>+</sup>, CD25<sup>-</sup>). While undifferentiated resting T cells were not affected, within the CD4<sup>+</sup>, CD25<sup>-</sup> cell population we observed effector-like and memory T cell with higher frequency (**Fig. 2.5 d and e**).



## Results (I)

---

Because increased TFHs and germinal center B cells which produce autoantibodies can be the premise for autoimmune diseases onset, we aged *Suv420h2 ko* mice for 1 year. Over this period, animals were periodically checked to monitor whether any immune related disorder arose. *Suv420h2 ko* mice did not develop any evident disease with aging, and analysis of major immune populations did not show evident anomalies (data not shown).

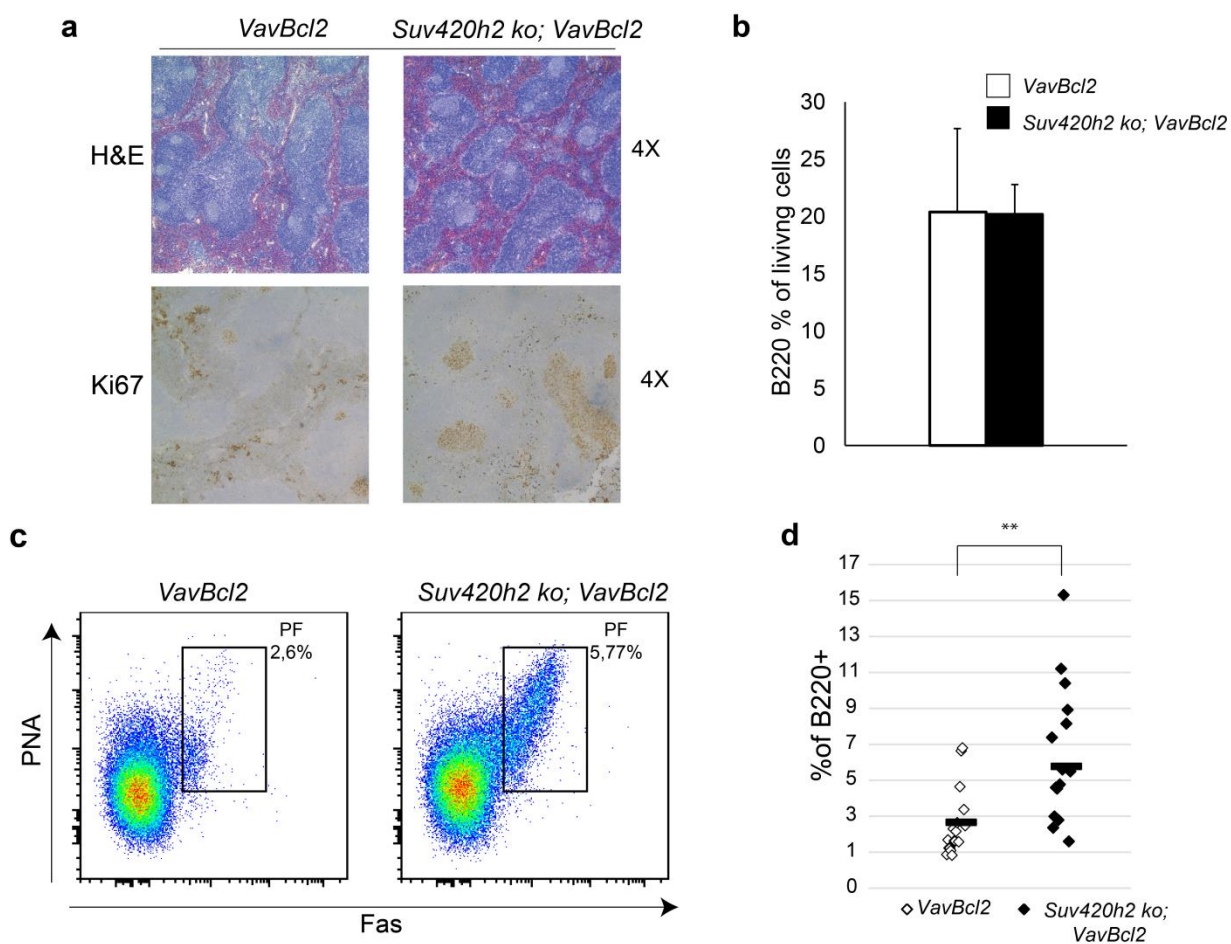
### **2.1.5 Release of apoptotic checkpoints exacerbates germinal center activation in *Suv420h2 ko* mice**

*Suv420h2 ko* mice showed spontaneous activation of B cells in the spleen. Even though this condition correlates with autoimmune disease or with inflammatory state, we detected no sign of disease during aging. This can be explained by different reasons: (1) the observed germinal center reaction was not strong enough to induce disease; (2) germinal center B cells were eliminated by apoptosis; (3) cells involved in regulation of immune response were able to keep the germinal center B cells under control; (4) a combination of all. These possibilities are all plausible, but most of them difficult to proof.

Because it is already known that germinal center B cells undergo massive physiological apoptosis upon T cell engagement, we assumed that extemporaneous germinal center activation in *Suv420h2 ko* mice did not result in a massive inflammation or autoimmune disease, because cells are cleared away via programmed cell death.

To understand whether releasing the stringency of the apoptotic pathway would exacerbate the germinal center activation in *Suv420h2 ko* mice, we crossed *Suv420h2 ko* with *VavBcl2* transgenic mice, generating *Suv420h2 ko; VavBcl2* mice.

*VavBcl2* transgenic mice express the antiapoptotic *Bcl2* under the control of *Vav* promoter, which is active throughout all haematopoiesis. These animals develop splenic germinal center hyperplasia, a phenotype which consists in a significant enlargement of the spleen and an unrestrained amplification of germinal center B cells. Germinal center hyperplasia breaks out in *VavBcl2* mice when they are 18 weeks old (Egle et al., 2004). Because at this stage germinal center are completely enlarged, any additional phenotype due to loss of *Suv420h2* in the *VavBcl2* background could have been masked by the fact that spleens are already too compromised. In this regard, we choose to analyse spleens from 8-10 weeks old mice. In average, *VavBcl2* spleens weighed 0.2 grams; 3 fold more than wild type spleens. When we weighted *Suv420h2 ko; VavBcl2* spleens, we appreciated that their size was increased by 1/3 compared to *VavBcl2* mice.



**Figure 2.6** *VavBcl2* mice show premature germinal center overexpansion in the absence of *Suv420h2*

**a)** Spleen sections from 8 weeks old mice were stained with haematoxylin/eosin to check spleen organization and ki67 to check proliferation. **b)** Percentage of splenic B220+ cell population checked by FACS. **c)** Flow cytometry plot depicting germinal center B cells (B220+, PNA+, FAS+). **d)** Statistics from **c)** (n=14). \*\*P < 0.01 (unpaired two-tailed Student's t-test).

To better investigate whether reduced apoptosis could contribute to amplify *Suv420h2 ko* germinal center B cells proliferation, we stained spleen sections with haematoxylin/eosin to check any significant change in organ architecture was detectable. Although *Suv420h2 ko VavBcl2* spleens maintained their structure, and all regions were still distinguishable, we noticed that all germinal centers were significantly enlarged and that they all contained highly proliferating cells; while *VavBcl2* showed only minimal signs of germinal center hyperplasia (**Fig. 2.6 a**). Then, we started investigating PNA+ FAS+ B cells in *Suv420h2 ko; VavBcl2* mice. Consistently with what we observed in the spleen sections, *VavBcl2* mice did not produce many PNA+ FAS+ B cells. *Suv420h2 ko; VavBcl2* mice, instead, showed a clear amplification of this cell population, although overall, splenic B cells were comparable to *VavBcl2* mice (**Fig. 2.6 b-d**).

## Results (I)

All together these data demonstrate that (1) loss of *Suv420h2* anticipates and exacerbates *VavBcl2* mice germinal center hyperplasia and that (2) apoptosis is one of the constraints which prevent full expansion of *Suv420h2 ko* germinal center B cells.

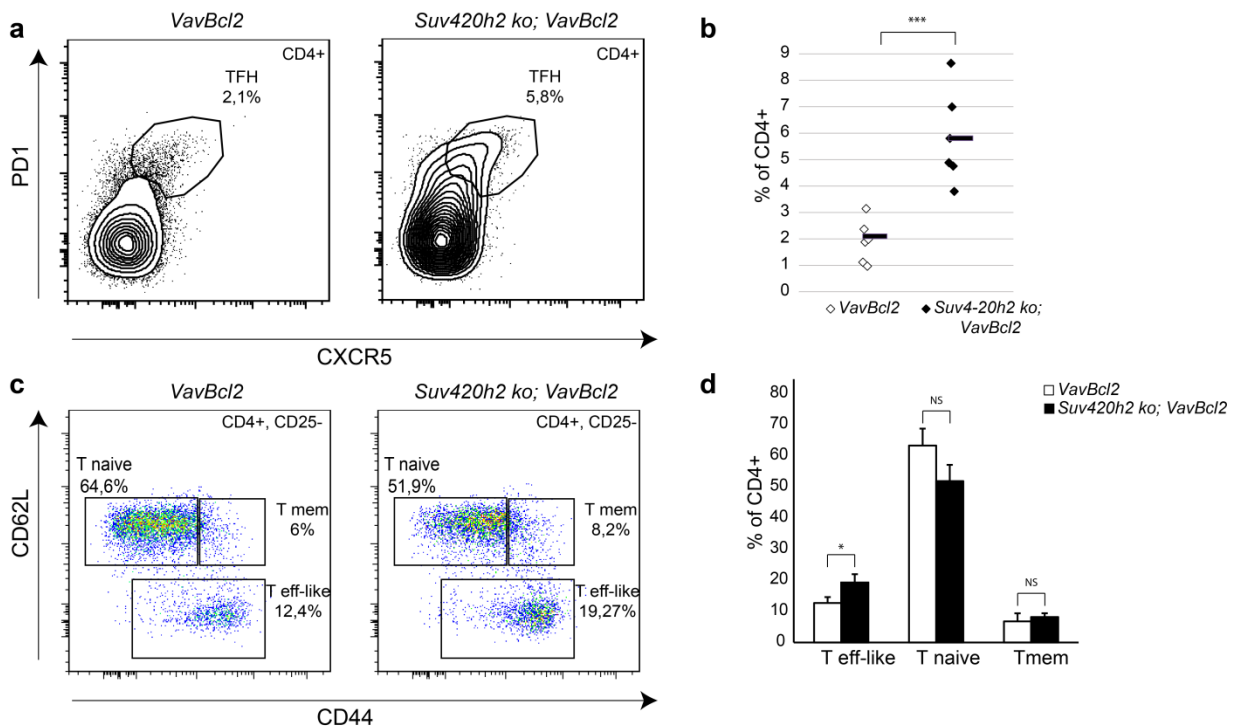
### 2.1.6 *Suv420h2 ko; VavBcl2* mice have increased proportion of follicular helper T cells

In previously published work, Egle *et al.* demonstrated that *Bcl2* overexpression under control of a specific B cell promoter does not result in spontaneous activation of the germinal centers. These data proved that excessive germinal center formation in the VavP-Bcl2 mice could not be simply a function of the amount of transgenic *Bcl2* in B cells, but rather other splenic populations are capable to induce uncontrolled B cell proliferation (Egle et al., 2004). In the same work, VavP-Bcl2 mice were crossed with GK5 mice, which lack CD4 cells. Like GK5 mice, bitransgenic VavP-Bcl2-GK5 mice had no CD4<sup>+</sup> T cells. When age-matched littermates were analysed at 18 weeks of age, it became clear that germinal centers were not detectable by histology. This finding suggested that expansion of the germinal centers in VavP-Bcl2 mice and the substantial increase in the frequency of activated B cells is dependent on CD4<sup>+</sup> T-cell help.

Because germinal center reaction is strictly connected to the presence of TFHs in the follicles we first investigated TFHs in the *VavBcl2* spleens. Surprisingly, *VavBcl2* TFHs were comparable, in percentage, to the wild type population (**Fig. 2.7 a**). This observation suggests that the signal which primed germinal center hyperplasia in *VavBcl2* mice did not come from TFHs; therefore another population of T helper cells must induce B cell proliferation. Because *Suv420h2 ko; VavBcl2* mice showed an increase in germinal center formation when compared to *VavBcl2*, we asked if these additional PNA<sup>+</sup> FAS<sup>+</sup> B cells were induced by TFHs. Remarkably, *Suv420h2 ko; VavBcl2* TFHs were consistently increased by 3 folds (**Fig. 2.7 a and b**).

*Suv420h2 ko; VavBcl2* bitransgenic mice represent another context where we could confirm, with better resolution, what we already observed in *Suv420h2 ko* mice. Moreover, these mice represent a very useful tool to perform in vitro analysis and additional molecular studies, as increased spleen cellularity will allow us to harvest enough material for further investigation. In fact, although *Suv420h2 ko* mice have increased TFHs and germinal center B cells, the total number of cells might be still limiting.

## Results (I)



**Figure 2.7 TFH cells are significantly expanded in *Suv420h2 ko; VavBcl2* spleen**

**a)** TFH cells (CD4+, CXCR5+, PD1+) population comparison between *VavBcl2* and *Suv420h2 ko; VavBcl2 ko* spleens. **b)** Frequencies of TFH cells shown as percentage of CD4+ cells (n=6). \*\*\*P < 0.001 (unpaired two-tailed Student's t-test). **c and d)** FACS plot and statistics (n=4) of resting naive T cells (CD4+, CD62L high, CD44- /low, CD25-) cells, T effector-like (CD4+, CD44+, CD62L low, CD25-) and memory T cells (CD4+, CD44+, CD62L+, CD25-) frequencies in *VavBcl2* and *Suv420h2 ko; VavBcl2* spleens. NS, not significant and \*P < 0.05 (unpaired two-tailed Student's t-test).

## 2.2 Discussion (I)

Antigen presentation to naive T cells by APCs triggers a number of differentiation events which initiate the immune response to pathogens. Germinal center formation is one of the most elaborated immune answers which results from antigen presentation and mediates production of highly specific antibodies. Failure to induce proper immune responses results in immunodeficiency; while, at the opposite, aberrant activation of the immune system is the primary source for inflammatory diseases, autoimmune disorders and lymphoma.

Although the signalling pathways which trigger and modulate the immune response are abundantly explained, we still lack some knowledge regarding the functional role that histone modifications exert in this context. An effort in this direction has been recently made by Beguelin and colleagues demonstrated that EZH2, the enzyme responsible for H3K27me3 establishment, is required for germinal center formation. In fact, it seems that the repressive H3K27me3 is necessary to keep silent genes involved in plasma cell differentiation, as their derepression antagonizes and bypasses the germinal center reaction (Beguelin et al., 2013).

Additional evidences which underlined the importance of repressive histone marks during mature T cell differentiation have been collected in the Almouzni laboratory. This group demonstrated that maintenance of the SUV39H1–H3K9me3–HP1 $\alpha$  pathway is critical to ensure TH2 cell stability by inhibiting TH1 cell-associated genes, indicating that H3K9me3 repressive mark is essential to silence inappropriate transcriptional programs (Allan et al., 2012).

The role of the repressive mark H4K20me3 during hematopoietic development has been already demonstrated by Schotta *et al.* In this paper, loss of *Suv420h1* and *Suv420h2* during haematopoiesis impaired lymphoid differentiation and resulted in ineffective B cell class switch recombination. Up to the present, this publication remains the only descriptive study which established an *in vivo* link between SUV420H-mediated H4K20me3 and hematopoietic development. Interestingly, in this thesis we described that the sole loss of *Suv420h2* did not impair lymphoid development, but instead resulted in spontaneous germinal center B cell activation. In homeostatic condition germinal center B cells are not present; therefore loss of *Suv420h2* and its associated H4K20me3 must generate positive signals to initiate the germinal center reaction. In this regard the work published by Stender and colleagues, encouraged us to believe that H4K20me3 represent a suppressive signal for several immune-related genes and its removal is required to initiate their expression. Since these results were obtained from *in vitro* derived macrophages and the histone methyltransferase which mediates H4K20me3 silencing of pro inflammatory genes is SMYD5 and not SUV420H2, we cannot decipher our

## Discussion (I)

---

phenotype in the light of these findings; however it is exciting to observe the presence of H4K20me3 outside its conventional heterochromatic location (Stender et al., 2012).

In the first place, we connected the production of *Suv420h2 ko* germinal center B cells mice to the B cell defects observed in *Suv420h* conditional knockout animals, implying a major role of SUV420h enzymes during B cell differentiation. For this reason we assumed that the observed defects were B cell intrinsic. However; *in vitro* experiments demonstrated that *Suv420h2 ko* B cells activated by different stimuli behaved comparably to wild type cells, suggesting that aberrant *Suv420h2 ko* B cell activation is not cell autonomous.

Making a step backwards in the germinal center formation timeline, we encounter a very specialized T helper cell population, known as TFH. TFHs differentiate from resting CD4<sup>+</sup> T cells and gradually spread into follicular areas where they start to interact with B cells, initiating germinal center reaction. Notably, increased germinal centers nicely correlated with presence of TFHs in the absence of *Suv420h2*. Furthermore, effector T cells were also increased in the absence of *Suv420h2* (**Fig. 2.5 d and e**), confirming signs of activation within CD4<sup>+</sup> T cell population. Recent work demonstrated that TFHs survive in the blood as memory T cells (Schmitt et al., 2014). Nicely, we found that also this population was increased compared to wild type mice, reinforcing the idea that *Suv420h2 ko* mice are able to differentiate more TFHs in the absence of immune stimuli.

Because germinal center B cells undergo massive apoptosis due to removal of potential self-reactive-antibody-producing cells, we verified whether releasing apoptotic check point had any impact on the observed phenotype. To do this we generated *Suv420h2 ko; VavBcl2* transgenic mice and checked how germinal center B and TFH cells behaved in young mice. Importantly, young *VavBcl2* animal's germinal centers were not significantly increased and TFHs were comparable to wild type mice. Differently, *Suv420h2* loss in the *VavBcl2* background showed a significant increase in the germinal center B cell population accompanied by a 3 fold increase of TFHs. These observations consistently confirmed what we already observed in *Suv420h2 ko* mice.

Because *Suv420h2 ko* mice are full knockout animals, we could not attribute the origin of the phenotype to any specific cell type, however we planned to perform bone marrow transplantation with *Suv420h2 ko; VavBcl2* B cells or naive T cells to reveal which of these populations is the actual inducer of spontaneous germinal center formation. Also, *Suv420h2 ko; VavBcl2* mice represent a promising tool to better investigate the molecular details of the described phenotype. In fact, we are going to test whether *in vitro* stimulation of *Suv420h2 ko; VavBcl2* B or/and T cells will show different behaviour compared to control cells.



## Discussion (I)

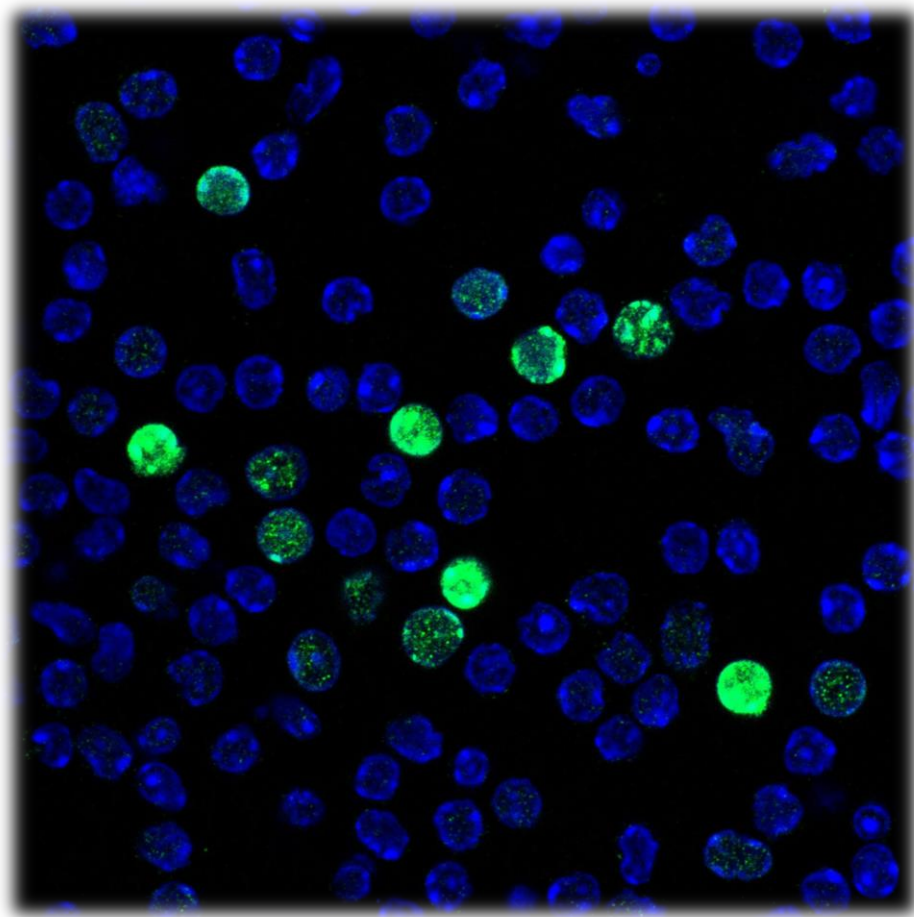
---

*VavBcl2* mice are prone to get follicular lymphoma after 1 year of age. Interestingly, preliminary data showed that *Suv420h2 ko; VavBcl2* mice die earlier compared to *VavBcl2* (data not shown). This phenotype needs to be further investigated, since we do not really know at the moment why loss of *Suv420h2* would further compromise *VavBcl2* mice viability. One hypothesis could be related to the onset of a more aggressive form of follicular lymphoma or a combination of tumour onset and persistent inflammation. To better address this question we set up an aging experiment where a cohort of mice is monitored for the appearance of any obvious disease.

In the near future we aim to identify genome wide H4K20me3 and SUV420H2 binding sites, in B and T cells. So far, ChIP-Seq analysis has been performed in our laboratory in MEF cells and revealed that SUV420H2-mediated H4K20me3 mainly covers retrotransposable elements and constitutive heterochromatic regions (telomeres, centromeres). However, the genomic distribution of this mark might change according to the cellular context we take in account. Once these target regions will be identified, we are going to analyse whether loss of H4K20me3 also results in gene activation which might promote terminal differentiation of mature immune cells.

Noteworthy, stacks of confocal images showed that only a very small percentage of CD4<sup>+</sup> T cells carried a clear H4K20me3 methylation pattern, while the others seemed to be totally refractory to the establishment of this repressive mark (**Fig. 2.8**). To better elucidate this point, we will sort CD4<sup>+</sup> T cell subpopulations to verify whether there is a preferential accumulation of H4K20me3 in a certain cell type. Moreover, we also noticed that the H4K20me3 signal not always overlapped with DAPI dense regions, suggesting that the H4K20me3 methylation pattern might be slightly different from the one observed in MEFs.

Although the limited amounts of data which have been collected so far leave little room for speculations and conclusions, we are pretty confident that H4K20me3 plays a role during germinal center formation. As yet, how this happens is unclear; however additional experiments will hopefully reveal more about the molecular basis of this phenotype.



**Figure 2.8 Heterogeneous H4K20me3 mark distribution within CD4+ T cell population**

a) Confocal picture showing CD4+ T cells stained with H4K20me3 primary antibody detected with Alexa488 secondary staining.



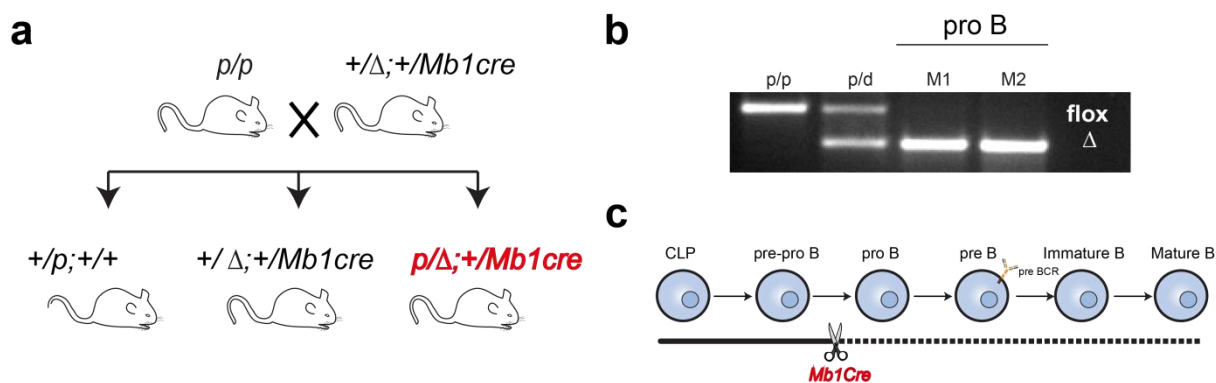
### 3 Results (II)

#### 3.1 SETDB1-mediated silencing of MuLVs is essential for B cell development

##### 3.1.1 Loss of *Setdb1* impairs B cell development

To investigate the role of *Setdb1* during B cells development, we generated *Setdb1<sup>p/Δ</sup>Mb1cre* mice, hereafter referred as *Setdb1<sup>Mb1</sup>* (**Fig. 3.1 a**). These conditional knockout mice (cKO) carried an engineered *Setdb1* floxed allele which loxP sites flanked exon 4; a *Setdb1* deleted allele and the *cre* recombinase. The expression of the recombinase was driven by the *Mb1* (CD79α) promoter, which starts to be active at early stages of B cell development. As soon as *Mb1cre* starts to be expressed, the *Setdb1<sup>fllox</sup>* allele is recognized and excised from the DNA, resulting in complete deletion of *Setdb1* starting from the pro B cell stage (**Fig. 3.1 b and c**).

B cells are produced and develop in the bone marrow through a series of well-defined differentiation stages and then relocate in peripheral organs. CLPs are the earliest B cells progenitors which gradually become pro B cells passing through the pre-pro B cell stage. Pro B cells start to rearrange DNA in order to produce the heavy immunoglobulin chain. This chain will be expressed on the pro B cells surface inducing a burst of proliferation which will direct cells to the pre B cell stage. At this point, light immunoglobulin chains will start to be rearranged to form functional immunoglobulins. Next, B cells will reach the immature B stage which is marked by surface expression of the immunoglobulins M (IgM). Some of these cells will become immature, immunoglobulin D (IgD) expressing B cells in the bones; while others will leave the bones to terminate their differentiation at other sites (**Fig. 3.1 c**). After maturation, the main organs which are populated by B cells are the spleen, the lymph nodes and the Peyer's patches. Anomalies in the size of these organs are already suggestive of aberrant B cell development.

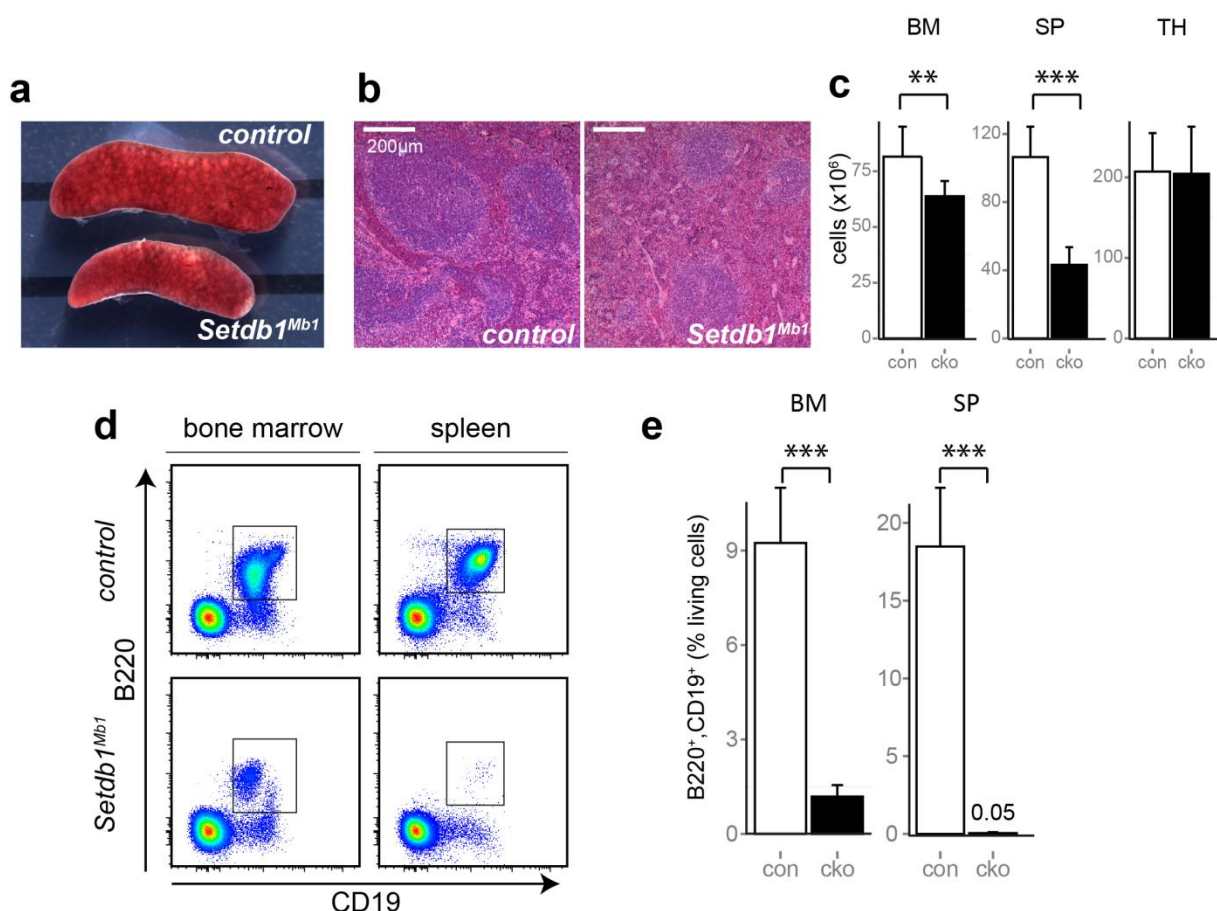


**Figure 3.1 Generation of *Setdb1*<sup>Mb1</sup> conditional knockout mice**

**a)** Breeding scheme adopted to obtain control and *Setdb1*<sup>Mb1</sup> cko mice. **b)** *Setdb1* deletion rate measured via PCR on genomic DNA from pro B cells sorted with FACS Aria. **c)** Graphic representation of B cell development and *Mb1cre* induced deletion.

Interestingly, the first macroscopic difference we could appreciate when we analysed *Setdb1*<sup>Mb1</sup> mice was a twofold reduction of the spleen size. This observation was followed by microscopic analysis of spleen architecture which revealed a significant decrease of the follicular areas, splenic regions where B cells typically collect (**Fig. 3.2 a and b**).

To investigate whether *Setdb1*<sup>Mb1</sup> B cells were impaired, bone marrow and splenic B cells were first stained with  $\alpha$ -CD19 and  $\alpha$ -B220 antibodies. CD19<sup>+</sup> B220<sup>+</sup> cell population was drastically reduced in *Setdb1*<sup>Mb1</sup> bone marrow and completely absent in the spleen, suggesting that loss of *Setdb1* significantly affects B cell development (**Fig. 3.2 d and e**).



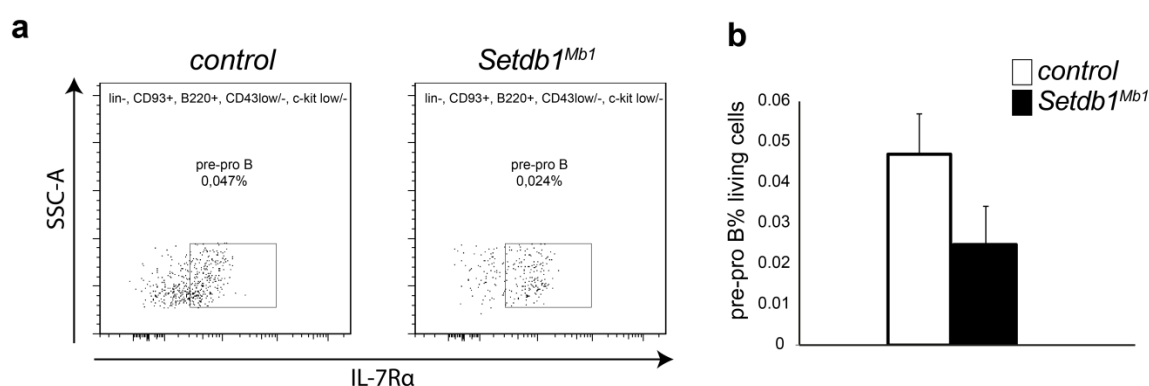
**Figure 3.2 *Setdb1*<sup>Mb1</sup> mice show severely reduced B cell population in the bone marrow and lack mature B cells**

**a)** Splenic from control and *Setdb1*<sup>Mb1</sup> mice. **b)** Spleen sections stained with haematoxylin/eosin. **c)** Cellularity of bone marrow (BM), spleen (SP) and thymus (TH) in control (con) and *Setdb1*<sup>Mb1</sup> (cko) mice (n=6). \*\*P < 0.01 and \*\*\*P < 0.001 (unpaired two-tailed Student's t-test). **d)** Flow cytometry of bone marrow and splenic B cells (B220<sup>+</sup>, CD19<sup>+</sup>). **e)** Statistics from d) shown as percentage of living cells (n=6). \*\*P < 0.01 and \*\*\*P < 0.001 (unpaired two-tailed Student's t-test).

## Results (II)

To better understand which stage of B cell differentiation was impaired in *Setdb1*<sup>Mb1</sup> mice, we analysed via multicolour flow cytometry all B cell compartments.

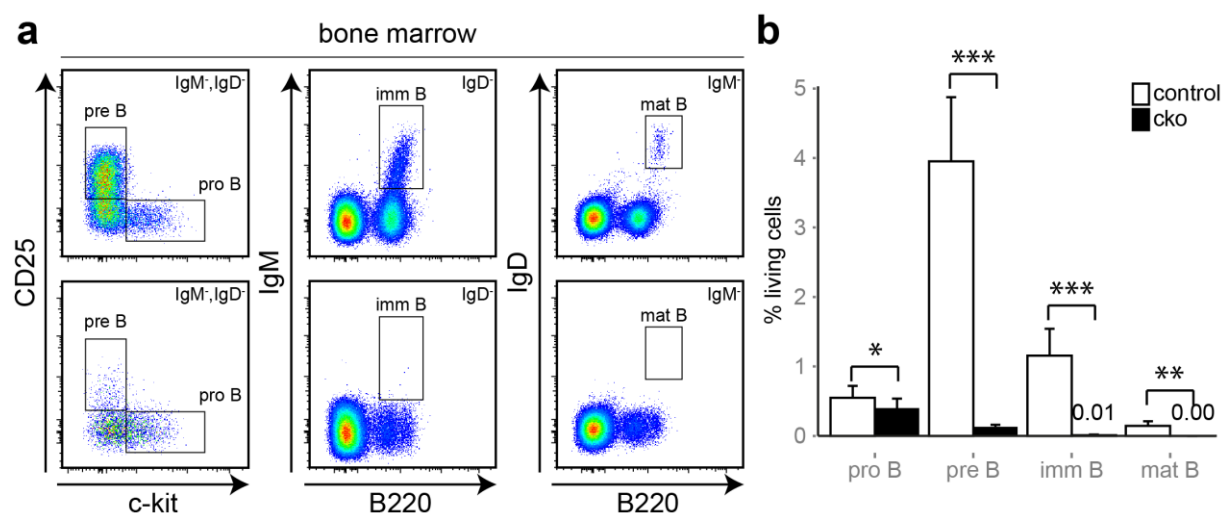
*Mb1* promoter reaches its peak of activation at pro B cell stage; however it starts to transcribe already in pre-pro B cells. *Setdb1*<sup>Mb1</sup> pre-pro B cells were reduced by nearly 2 fold (**Fig. 3.3 a and b**); nevertheless they could still generate a largely normal pro B cell population, indicating that partial expression of the cre recombinase did not irreversibly affect pre-pro B to pro B cell transition (**Fig. 3.4 a left panel**).



**Figure 3.3** *Setdb1*<sup>Mb1</sup> pre-pro B cells are reduced

**a)** Flow cytometry analysis of pre-pro B cells (Lin-, CD93+, B220+, CD43-/low; c-kit-/low, IL-7Rα+).  
**b)** Statistics from a) shown as percentage of living cells (n=3).

Pro B stage represent a critical step during B cell development, because at this point the heavy chain of the immunoglobulin starts to rearrange, addressing cells to a permanent B cell fate. Successful rearrangements result in the expression of the pre-BCR which induces a burst of cell proliferation, marking the entrance into the pre B stage.



## Results (II)

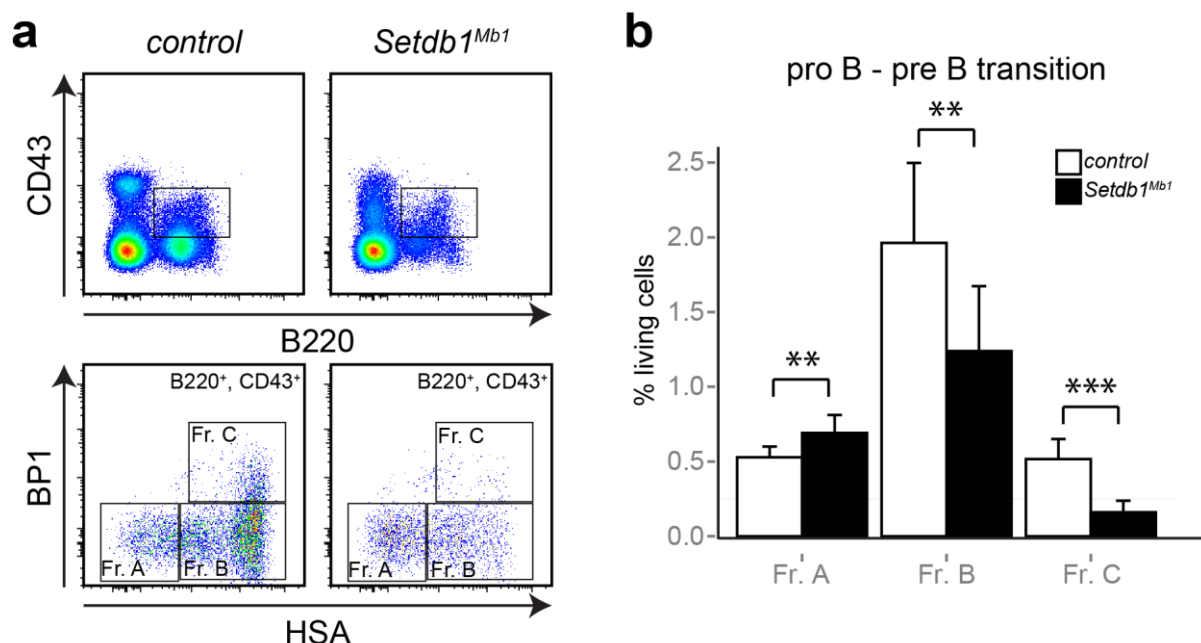
**Figure 3.4 Impaired pro B to pre B cell transition in *Setdb1*<sup>Mb1</sup> mice**

**a)** FACS analysis of the main B cell populations in the bone marrow: pro B (CD19+, IgM-, IgD-, CD25-, c-kit+), pre B (CD19+, IgM-, IgD-, CD25+, c-kit-), immature B (B220+, IgD-, IgM+), mature B (B220+, IgD+, IgM-). **b)** Statistics from a) shown as percentage of living cells (n=6).

\*P < 0.05, \*\*P < 0.01, \*\*\*P < 0.001 (unpaired two-tailed Student's t-test).

As shown in **Fig. 3.4** *Setdb1*<sup>Mb1</sup> mice generated normal pro B cells numbers; however only few of them passed the checkpoint to the pre B stage, as they were reduced by 30 fold compared to control mice. Moreover, *Setdb1*<sup>Mb1</sup> pre B cells did not proceed to subsequent step of B cell commitment, demonstrating that loss of *Setdb1* sharply abrogates B cell development at the pre B cell stage (**Fig. 3.4 a and b**).

To look with better resolution at the pro B to pre B cell switch in *Setdb1*<sup>Mb1</sup> mice we used Hardy's marker scheme (CD43, B220, HSA/CD24, BP1) for B cell fractionation (Hardy et al., 1991).

**Figure 3.5 Loss of *Setdb1*<sup>Mb1</sup> B cell occurs at late pro B cell stage**

**a)** FACS plots showing pro B to pre B cell transition according to the Hardy scheme. B220+, CD43+ population encompasses early and late pro B cells together with small pre B cells. These 3 subpopulations are discriminated in fraction A, B and C according to the expression level of HSA and BP1 markers. **b)** Statistics from a) shown as percentage of living cells (n=6). \*\*P < 0.01 and \*\*\*P < 0.001 (unpaired two-tailed Student's t-test).

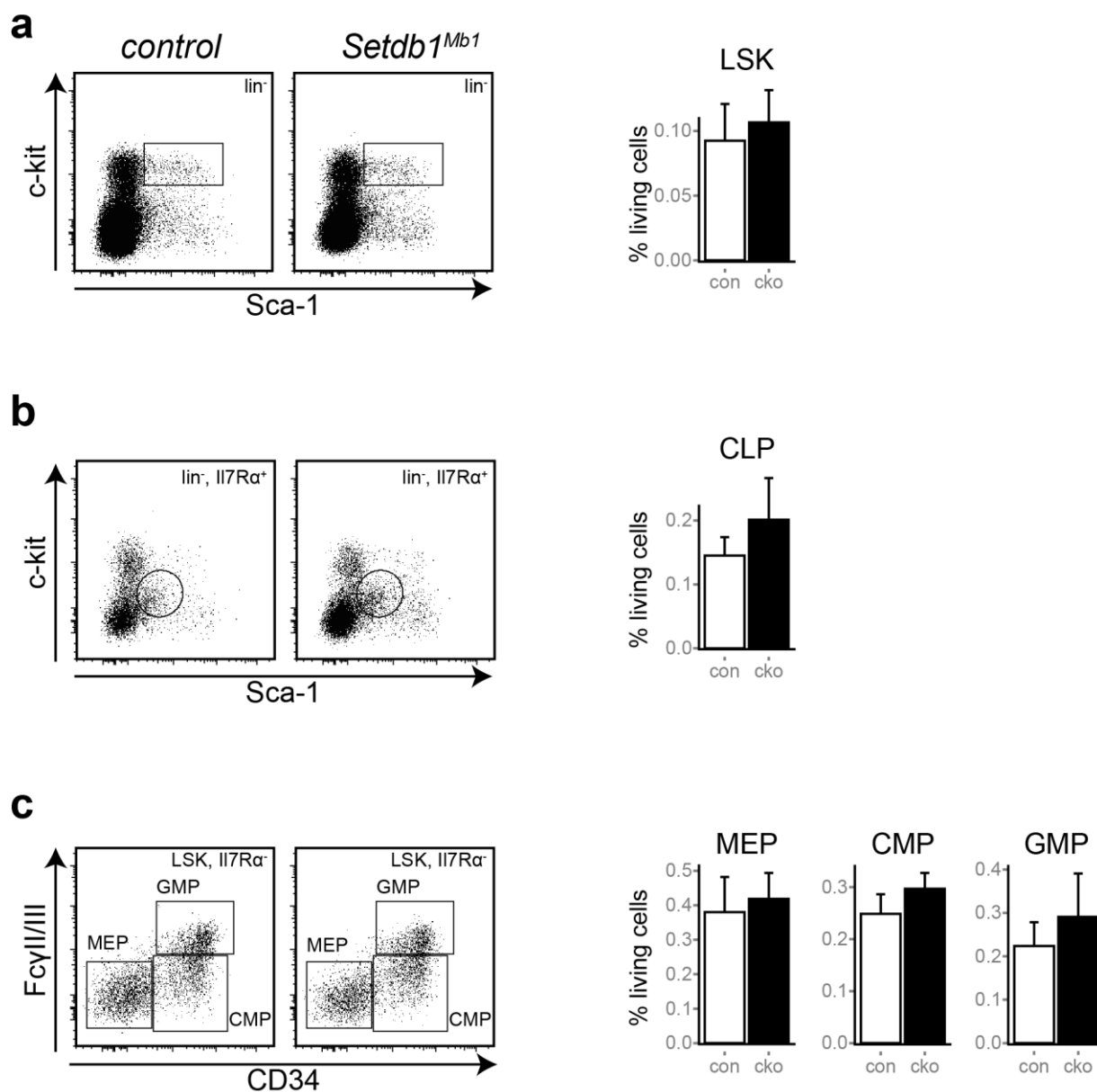
This marker combination allows discriminating different populations within the CD43+ B220+ cells: fraction A (pre-pro B and early pro B; HSA low/+ BP1-), fraction B (late pro B; HSA+/+ BP1-) and fraction C (large pre B; HSA++ BP1+). Coherently with the data shown above, early pro B cells were not affected in the absence of *Setdb1* (Fr.A). However, as soon

## Results (II)

as B cells progressed towards the late pro B cell stage, they started to decrease (Fr.B) till disappearing in the large pre B cell fraction (Fr.C) (**Fig. 3.5 a and b**).

To exclude that the observed B cell defects could be caused by ectopic expression of the cre recombinase, we also analysed hematopoietic progenitor populations. Importantly, all hematopoietic precursors were not affected in *Setdb1*<sup>Mb1</sup> mice (**Fig. 3.6 a-c**).

Collectively, these data demonstrated that *Setdb1* is essential for B cell development.

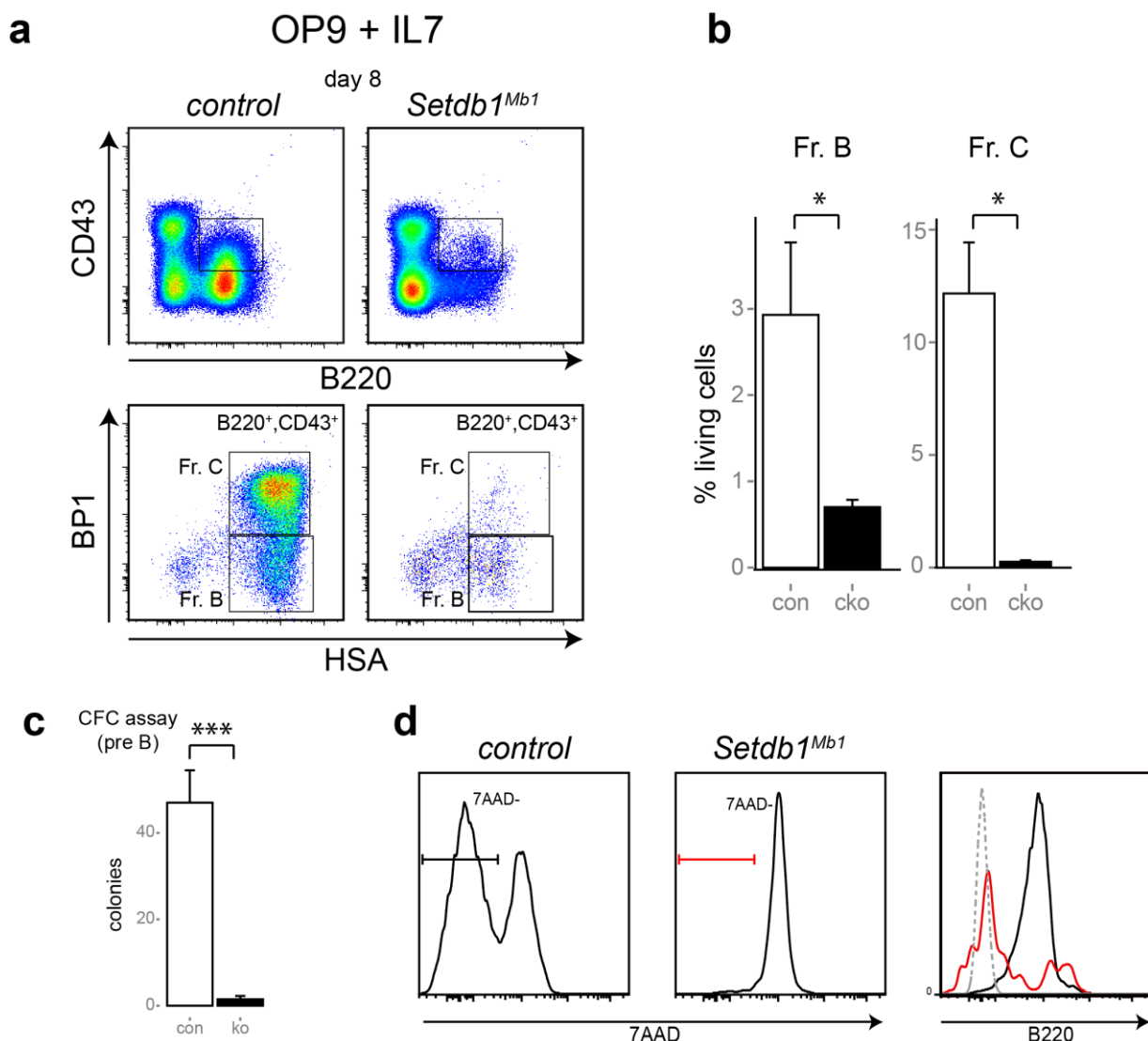


**Figure 3.6 Normal frequencies of *Setdb1*<sup>Mb1</sup> hematopoietic progenitors**

**a**) LSKs (Lin<sup>-</sup>, Sca<sup>+</sup>, c-kit<sup>+</sup>) hematopoietic stem cells. **b**) CLPs (Lin<sup>-</sup>, IL7Rα<sup>+</sup>, Sca low, c-kit low) and **c**) myeloid progenitors: CMPs (Lin<sup>-</sup>; IL7Rα<sup>-</sup>; c-kit high, Sca<sup>-</sup>; CD34<sup>+</sup> and CD16/32<sup>-</sup>); GMPs (Lin<sup>-</sup>; IL7Rα<sup>-</sup>; c-kit high, Sca<sup>-</sup>; CD34<sup>+</sup> and CD16/32<sup>+</sup>) and MEPs (Lin<sup>-</sup>; IL7Rα<sup>-</sup>; c-kit high, Sca<sup>-</sup>; CD34<sup>-</sup> and CD16/32<sup>-</sup>). To each FACS plot are associated statistics shown as percentage of living cells (n=5).

### 3.1.2 Intrinsic role of *Setdb1* during B cell development

To test whether B cell developmental arrest in *Setdb1*<sup>Mb1</sup> mice was due to B cell intrinsic defects, we first performed *in vitro* pre B cell colony assay using MethoCult M3630 supplemented with IL-7. As shown in **Fig. 3.7 c** after 10 days of culture, *Setdb1*<sup>Mb1</sup> bone marrow failed to generate pre B cell colonies. Additionally, while control cells were mostly negative to the viability dye 7AAD and were expressing high level of B220 and CD19; nearly all *Setdb1*<sup>Mb1</sup> cells were positive to 7AAD and B cell markers were hardly expressed on their surface (**Fig. 3.7 d**). Because *Setdb1*<sup>Mb1</sup> B cells are already compromised and are overall underrepresented in the bone marrow, we repeated the assay using MACS sorted progenitor cells where *Setdb1* deletion hasn't occurred yet. Even in this case, *Setdb1*<sup>Mb1</sup> progenitors did not form any colony (data not shown).



**Figure 3.7 Intrinsic role of *Setdb1* during B cell development**

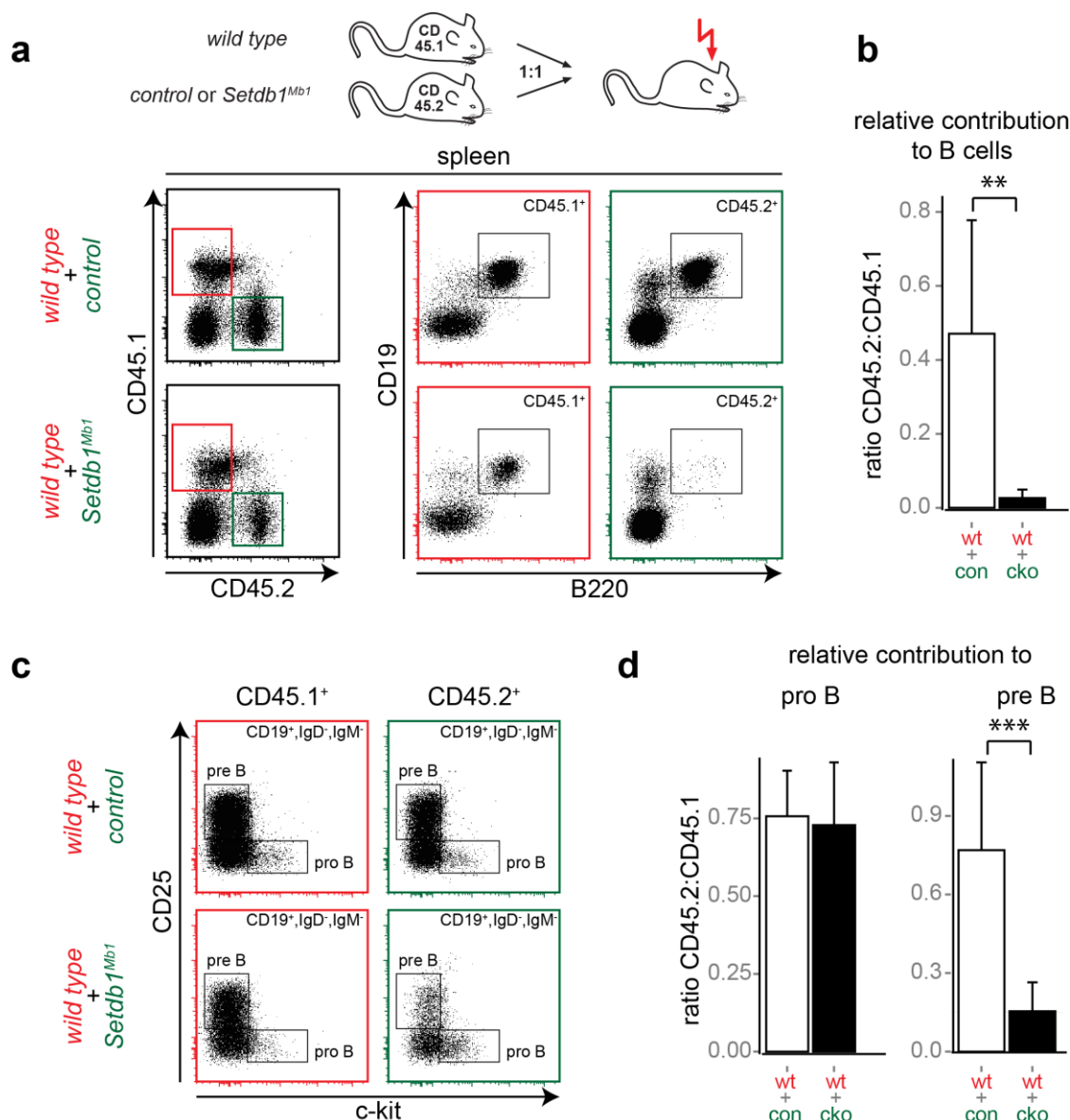
**a)** Pre B cell differentiation assay. Lineage negative progenitors were co-cultured with OP9 stromal cells in the presence of IL-7 for 8 days and then analysed by FACS using Hardy's scheme. **b)** Statistics



## Results (II)

from a). \* $P < 0.05$  (unpaired two-tailed Student's t-test). **c**) pre B cell colony assay. Control and *Setdb1*<sup>Mb1</sup> bone marrow cells were grown on methylcellulose medium for 10-14 days. Colony number represents the average of 3 biological replicates. \*\*\* $P < 0.001$  (unpaired two-tailed Student's t-test). **d**) Colonies were harvested after 10 days of culture and analysed by FACS for viability (7AAD) and expression of B cell markers. On the right FACS histogram shows B220 surface expression on *control* (black line) and *Setdb1*<sup>Mb1</sup> (red line) cells, both gated on 7AAD<sup>-</sup> fraction.

Additionally, we also set up OP9 coculture which allows pre B cell differentiation. Progenitor cells were seeded on OP9 stromal cells in the presence of IL-7 and the coculture was refreshed every other day. After 8 days cells were analysed by FACS and their ability to differentiate in pre B cells was monitored using the Hardy's scheme. Control CD43<sup>+</sup> B220<sup>+</sup> population distributed 80% of the cells in Fr.C and the 20% in Fr.B. *Setdb1*<sup>Mb1</sup> progenitors, instead, exhibited a drastic reduction of CD43<sup>+</sup> B220<sup>+</sup> cells which contained few pro B cells (Fr.B) and nearly no pre B cells (Fr.C) (**Fig. 3.7 a and b**).



**Figure 3.8 Intrinsic role of *Setdb1* during B cell development**

**a)** Competitive bone marrow transplantation. Bone marrow cells derived from 5-FU injected mice were mixed 1:1 and transplanted in sublethally irradiated mice.  $1 \times 10^6$  wild type (CD45.1) donor cells were mixed with equal amount of control or *Setdb1*<sup>Mb1</sup> competitor cells (CD45.2). Mice were analysed 6-9 weeks after transplantation. **b)** Statistics showing the relative contribution to the B cell populations depicted in a). \*\*P < 0.01 (unpaired two-tailed Student's t-test). **c)** pro B to pre B cell transition in transplanted mice. **d)** Statistics showing the relative contribution to pro B and pre B cell stages depicted in c). Bargraphs depict average and standard deviation from at least 5 biological replicates. \*\*\*P < 0.001 (unpaired two-tailed Student's t-test).

Transplantation experiments were performed in collaboration with Klein laboratory (Institute of Immunology, LMU Munich)

All these *in vitro* experiments already provided strong evidence that *Setdb1* loss intrinsically affect B cell development; however we also performed additional analysis to confirm *in vivo* the cell intrinsic nature of the observed phenotype. To address this point, competitive bone marrow transplantations were performed using 1:1 mixture of wild type (CD45.1) and *Setdb1*<sup>Mb1</sup> (CD45.2) progenitors or wild type (CD45.1) and *Mb1cre* (CD45.2) progenitors. 6-9 weeks after transplantation mice were analysed to check B cell development. *Mb1cre* transgenic bone marrow showed a largely normal B cell development, producing pro B and pre B cells in a similar fashion to the cotransplanted wild type cells. Then we tested B cell population in *Setdb1*<sup>Mb1</sup> transplanted animals. Consistently with what happens in non-transplanted mice, *Setdb1*<sup>Mb1</sup> donor cells produced significantly less amount of B cells in the bone marrow and showed almost complete absence of splenic B cells (**Fig. 3.8 a and b**). Analysis of B cells differentiation in the bone marrow displayed that pro B cell population were present with normal frequencies, while *Setdb1*<sup>Mb1</sup> pre B cell population was compromised (**Fig. 3.8 c and d**). To demonstrate that other lineages were not impaired, *Setdb1*<sup>Mb1</sup> transplanted cells were also checked myeloid cells. Importantly, Gr<sup>+</sup> Mac<sup>+</sup> population resulted to be comparable to the cotransplanted wild type cells (data not shown). Taken together these data demonstrate that *Setdb1* controls B cell development in a cell-intrinsic manner.

**3.1.3 *Setdb1*<sup>Mb1</sup> B cells ectopically transcribe murine retrotransposable elements**

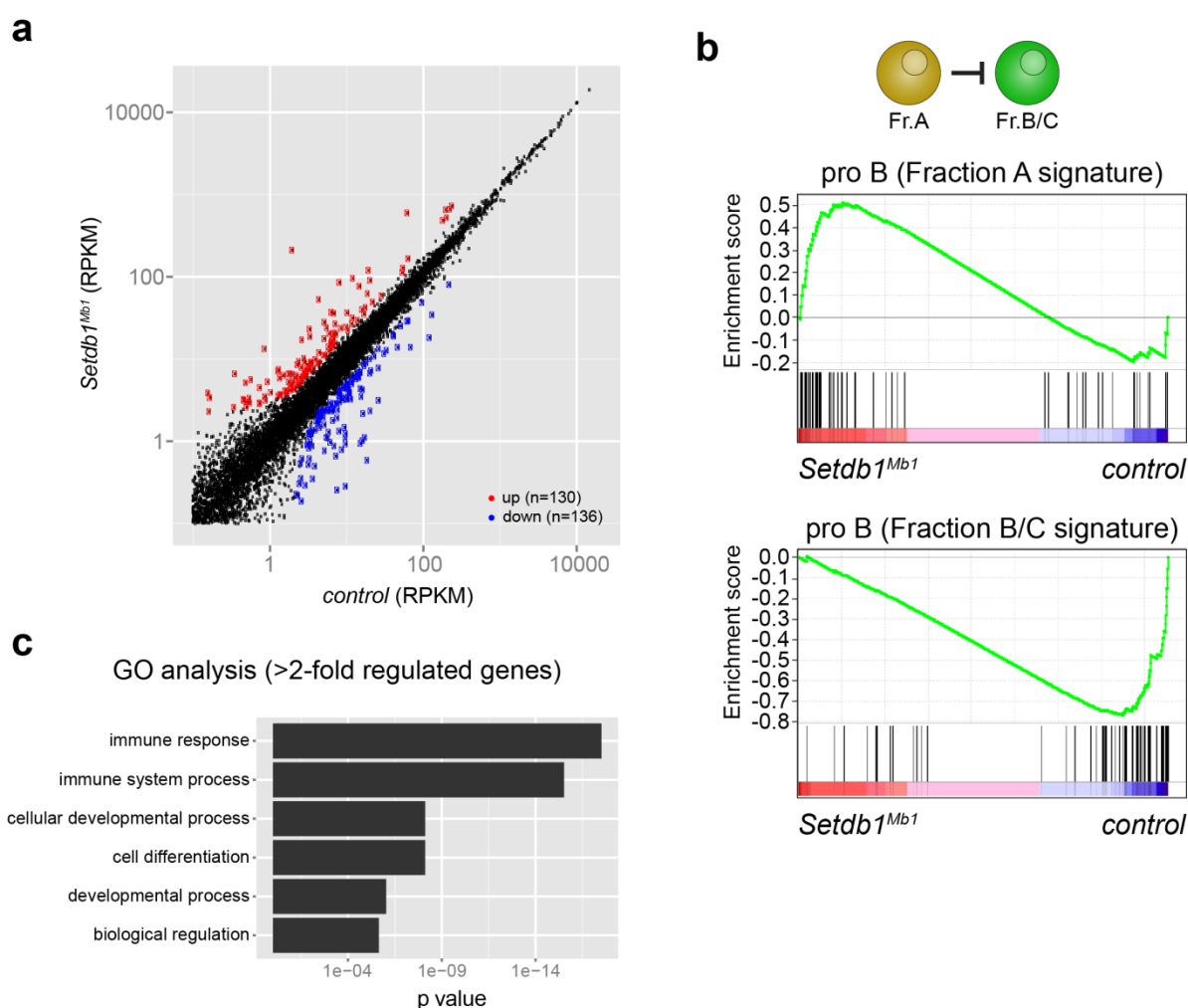
To learn more about the molecular mechanism which cause B cell failure in the absence of *Setdb1*, we performed RNA-Seq analysis on sorted pro B cells. Overall, we found that 260 genes were deregulated more than 2 folds in *Setdb1*<sup>Mb1</sup> pro B cells (**Fig. 3.9 a**). Considering the repressive function which has been continuously attributed to *Setdb1* in the last years, we expected to observe a clear bias towards gene upregulation. Surprisingly, within the cohort of deregulated genes, the number of down and upregulated transcripts was nearly the same,



## Results (II)

suggesting that loss of *Setdb1* does not result mainly in gene derepression (**Fig. 3.9 a; Appendix, Table 5.2**).

To reveal whether the group of deregulated genes was part of a specific regulatory pathway, we performed gene ontology (GO) analysis with the intent to group them according to their function and/or the biological processes they are involved in. As a result, we found that *Setdb1*<sup>Mb1</sup> pro B cells deregulated genes involved in immune regulation, as well as developmental and differentiation processes, reflecting the strong developmental block we observed during B cell differentiation (**Fig. 3.9 c**). Although GO analysis nicely confirmed that immune regulation was impaired; it was really challenging to guess which genes were directly responsible for B cell developmental arrest in *Setdb1*<sup>Mb1</sup> pro B cells, as none of them was a master regulator of B cell development.



**Figure 3.9** Gene expression analyses in *Setdb1*<sup>Mb1</sup> pro B cells

**a)** RNA-Seq performed on sorted pro B cells. Dot plot showing the number of up (red) and down (blue) regulated genes. **b)** Gene set enrichment analysis performed over all deregulated genes. *Setdb1*<sup>Mb1</sup> B cells exhibited enrichment for early pro B cells signature (Fr.A) and poorly expressed late pro B cells transcripts (Fr.B/C). **c)** GO analysis of >2 fold deregulated genes in *Setdb1*<sup>Mb1</sup> pro B cells.

In the first place, we carefully examined the function of all deregulated genes found in our list and checked whether they could be connected to the pro B cell developmental block. Analysis of the most deregulated transcripts revealed that any of them could be directly responsible for the B cell developmental failure since (1) many of these genes were already transcribed at such low level in control mice that further downregulation could not have any impact on transcriptional networks, (2) full or conditional knockout mouse model have been previously described for some of the analysed genes, but none of them showed B cell developmental arrest at the pro B cells stage, (3) genes modestly upregulated, for which no mouse model was available, were poorly transcribed in wild type pro B cells and (4) no B cell master regulator was found to be transcriptionally altered. All this led us to think that maybe gene deregulation was not the cause but rather the consequence of B cell failure upon *Setdb1* loss. We then compared all *Setdb1*<sup>Mbl</sup> pro B cells deregulated genes with the wild type Hardy's Fr.A and Fr.B signatures available at immgen.org. Notably, we found that *Setdb1*<sup>Mbl</sup> pro B cells accumulated early pro B cell transcripts, while late pro B cells related genes were poorly represented compared to control cells (**Fig. 3.9 b**), possibly reflecting the more immature stage already evident in the Hardy staining. Indeed, RNA-Seq was performed on sorted pro B cells (CD19+, IgM-, IgD-, CD25-, c-kit+) which corresponded to Hardy's Fr.A and Fr.B. Within these fractions we could notice a clear enrichment of *Setdb1*<sup>Mbl</sup> cells in Fr.A, while Fr.B contained significantly less events compared to control (**Fig. 3.5 a**). This bias was indicative for an evident retention of *Setdb1* deficient pro B cells in a more immature stage, as confirmed by the gene expression profile. All these observations suggested that gene deregulation might not be the direct cause for B cell failure and that some other events must interfere with the establishment of the B cell program in the absence of *Setdb1*.

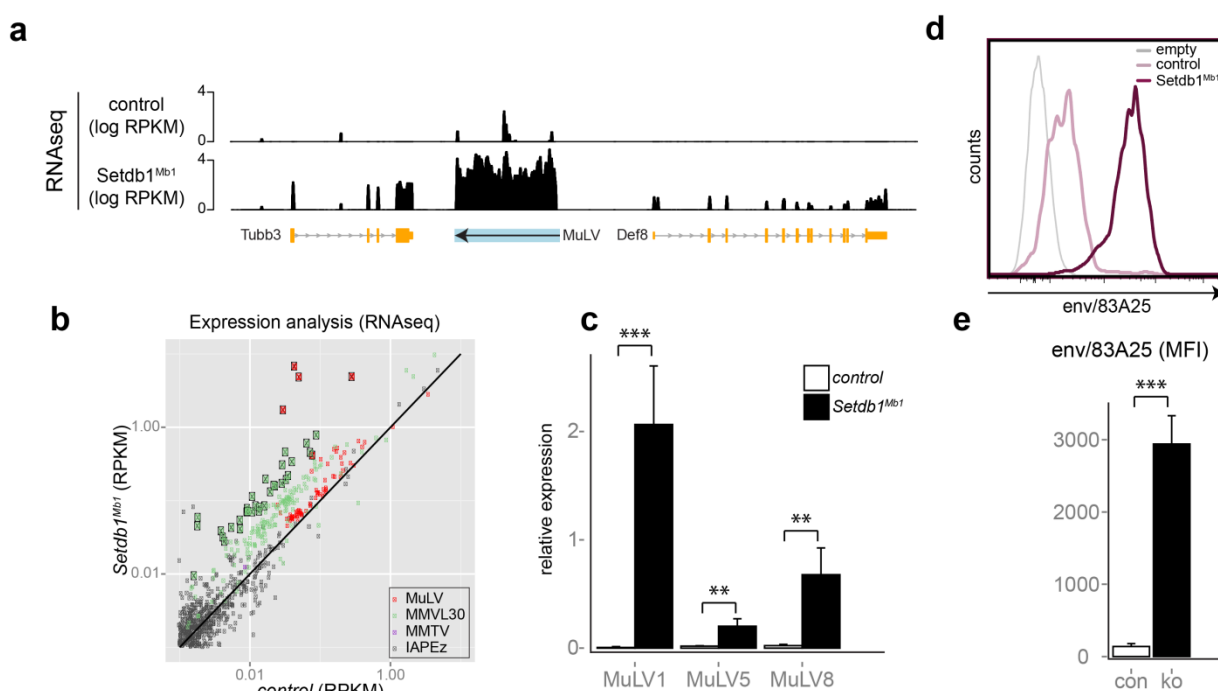
Recent literature demonstrated that SETDB1 induces retrotransposons silencing via establishment of the repressive mark H3K9me3 (Karimi et al., 2011). Because analysis of deregulated coding transcripts seemed not to be the direct cause of the phenotype, we started investigating the non-coding part the genome. At first, we checked whether non-coding elements were expressed in *Setdb1*<sup>Mbl</sup> pro B cells. Interestingly, among all families of endogenous retroviruses, two classes of retrotransposons were prominently transcribed. These elements are known as murine leukaemia viruses (MuLVs) and mouse VL30 retroelements (MMVL30) (**Fig. 3.10 a and b; Appendix, Table 5.3**), and they are able to cause cancer when ectopically expressed in the host (Herquel et al., 2013; Young et al., 2012).

To substantiate the transcriptional changes observed in our RNA-Seq profile, we designed a set of primers which amplified several regions of the MuLV elements and performed qRT-

## Results (II)

PCR on pro B cells cDNA derived from 6 biological replicates. Importantly, MuLV transcripts were not detectable in control pro B cells, while they were abundantly expressed in the absence of *Setdb1* (**Fig. 3.10 c**). Interestingly, MuLVs regions were not only transcribed but also translated, producing viral proteins. Bone marrow cells were incubated with 83A25, an antibody which recognizes viral envelope protein (MLV SU). 83A25 was then secondary labelled with Alexa647 to test its expression on the pro B cell surface. Consistently with the high level of retroviral transcripts, we also detected significant amount of the envelope viral protein on *Setdb1*<sup>Mb1</sup> pro B cells surface (**Fig. 3.10 d and e**).

Identification of MuLVs derepression, confirmed the recent findings which described SETDB1 as main silencer of endogenous retroviruses; however so far nobody has ever provided a comprehensive explanation of how overexpression of retrotransposons might interfere with developmental processes.



**Figure 3.10 Ectopic expression of retrotransposons in *Setdb1*<sup>Mb1</sup> pro B cells**

**a)** Number of reads across one representative MuLV elements in control and *Setdb1*<sup>Mb1</sup> pro B cells. **b)** Dot plot showing expression of endogenous retroviruses in control and *Setdb1*<sup>Mb1</sup> pro B cells. **c)** MuLVs transcription confirmed via q-RT-PCR (n=6). Several primer pairs were used to detect transcription of MuLV genes (gag, pol and env). Here we show transcription level of the gag transcripts from the highest upregulated MuLVs. \*\*P < 0.01 and \*\*\*P < 0.001 (unpaired two-tailed Student's t-test). **d)** Expression of envelop viral protein on the pro B cells surface detected by FACS analysis. Bone marrow cells were incubated with α-env (83A25) antibody followed by secondary detection with Alexa647. **e)** Bar graph showing env protein expression level calculated as average of the mean fluorescence intensity (n=6). \*\*\*P < 0.001 (unpaired two-tailed Student's t-test).

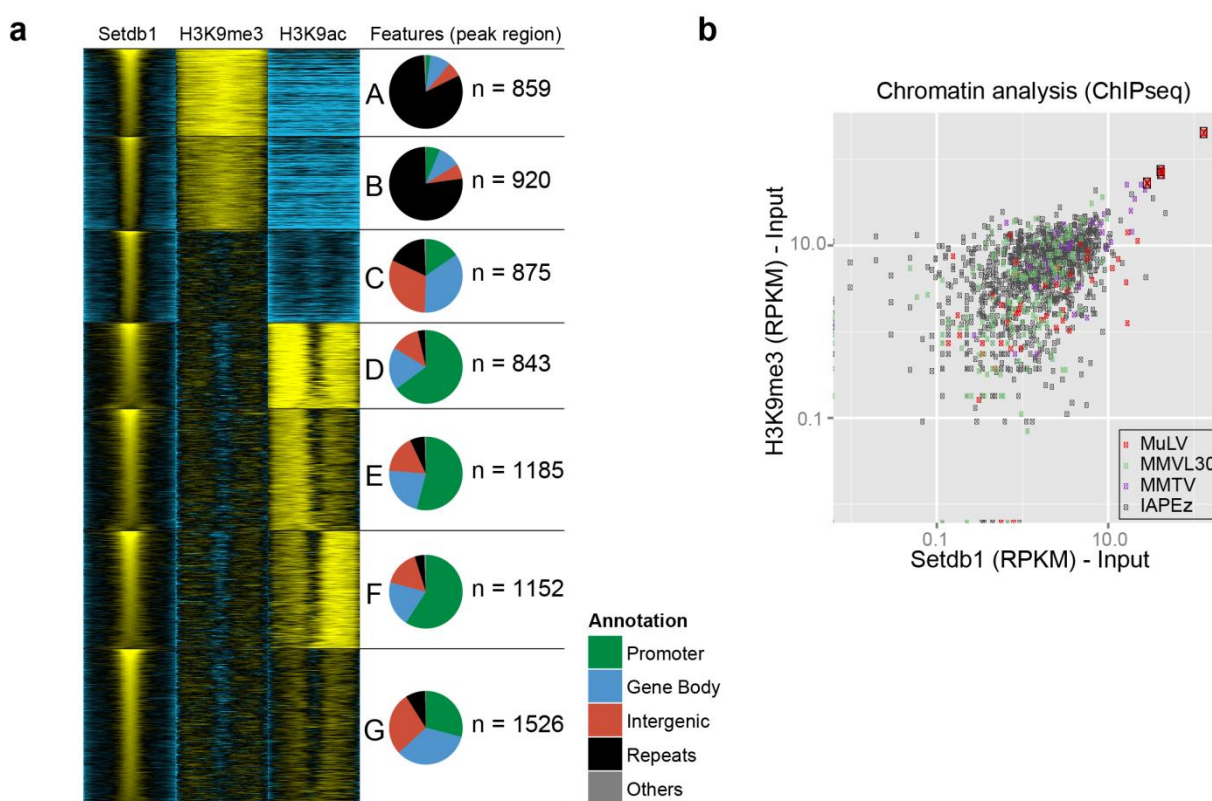
In summary, we conclude that aberrant gene expression in *Setdb1*<sup>Mb1</sup> pro B cells is not a direct consequence of *Setdb1* depletion, but rather results from other events which managed to retain

## Results (II)

cells in the early pro B stage. Because gene deregulation could not explain B cells developmental arrest in *Setdb1*<sup>Mb1</sup> mice; we propose that massive transcription of endogenous retroviral elements can be a direct consequence of *Setdb1* depletion. To corroborate this hypothesis we next performed ChIP-Seq experiments.

### 3.1.4 SETDB1 represses retrotransposons in pro B cells

To learn more about SETDB1 binding sites and to understand whether MuLVs silencing was due to SETDB1 mediated H3K9me3, we performed ChIP-Seq using short-term cultured *Rag2ko* pro-B cells. Chromatin from these cells was immunoprecipitated using  $\alpha$ -SETDB1,  $\alpha$ -H3K9me3 and  $\alpha$ -H3K9ac antibodies and analysed via high-throughput sequencing.



**Figure 3.11 SETDB1 and H3K9me3 enrichment across repetitive elements.**

**a)** ChIP-Seq analysis of *in vitro* expanded pro B cells performed with  $\alpha$ -SETDB1  $\alpha$ -H3K9me3 and  $\alpha$ -H3K9ac antibodies. This experiment was performed in collaboration with Busslinger laboratory (IMP, Vienna). **b)** Correlation analysis of SETDB1 and H3K9me3 across retrotransposable elements. The highest association was found over MuLV elements. **c)** Dot plot showing expression of endogenous retroviruses in control and *Setdb1*<sup>Mb1</sup> pro B cells.

Bioinformatic analysis of SETDB1 and H3K9me3 peaks showed that SETDB1 covered many more regions in the genome compared to H3K9me3. Unexpectedly, SETDB1 binding sites were rarely overlapping with H3K9me3 peaks, suggesting that this histone methyltransferase does not establishes H3K9me3 at all its binding sites.

## Results (II)

Subsequently, we started investigating SETDB1, H3K9me3 and H3K9ac binding sites genome-wide. SETDB1 and H3K9me3 peak colocalization excluded the presence of the active mark H3K9ac and, vice versa, H3K9me3 repressive mark was missing at loci where SETDB1 and H3K9ac gathered (**Fig. 3.11 a**). Next, we tried to unravel which genomic areas underlaid SETDB1/H3K9me3 and SETDB1/H3K9ac binding sites. Notably, SETDB1/H3K9me3 mostly covered repetitive elements, while SETDB1/H3K9ac peaks were found at promoter regions (**Fig. 3.11 a**).

Because MuLVs and MMLV30 expression in pro B cells might be a direct consequence of *Setdb1* depletion, we checked SETDB1 coverage across these elements. Remarkably, we found that SETDB1 clearly occupied these regions together with H3K9me3; strongly supporting the idea that ectopic transcription of these elements in pro B cells is directly connected to *Setdb1* loss (**Fig. 3.11 b**).

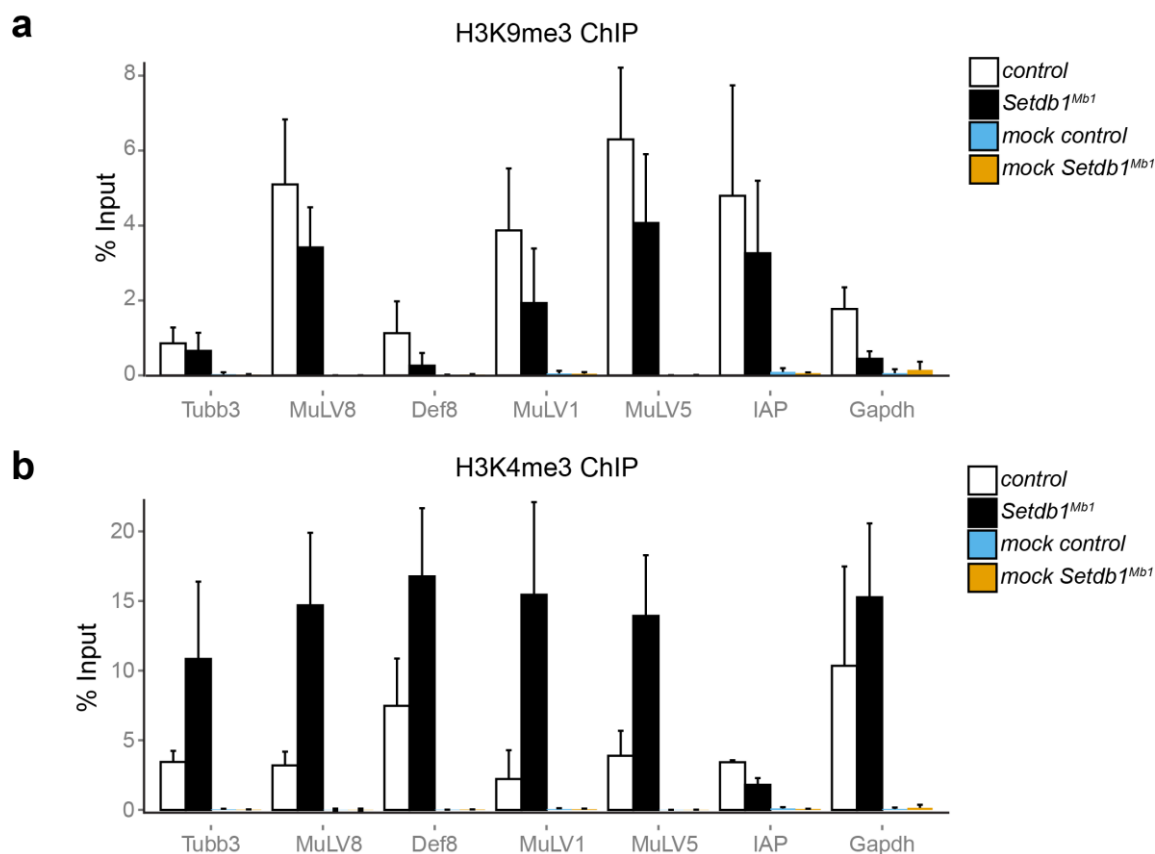
To test how repressive chromatin behaved upon *Setdb1* depletion, we sorted CD43<sup>+</sup> CD19<sup>+</sup> cells from control and *Setdb1*<sup>Mb1</sup> bone marrow and performed ChIP-qPCR for the repressive H3K9me3 and the active H3K4me3 chromatin marks. Because MuLV elements showed the highest transcription levels (**Fig. 3.10 b**), we choose these regions to test whether the distribution of the histone marks was changed in upon *Setdb1* depletion. H3K9me3 broadly covered MuLV elements in control cells, confirming what we observed in our ChIP-Seq profile. Unexpectedly, sorted *Setdb1*<sup>Mb1</sup> pro B cells displayed a reduction but not a complete loss of H3K9me3, indicating that MuLVs can be activated even if the repressive mark H3K9me3 is still present on the sequence (**Fig. 3.12 a**).

Because MuLVs are massively transcribed in *Setdb1*<sup>Mb1</sup> pro B cells we also performed ChIP-qPCR on sorted cells to test whether they gained H4K3me3 across their sequences. *Setdb1*<sup>Mb1</sup> cells showed a significant increase of H3K4me3 across MuLVs compared to control mice, which showed almost no enrichment of this mark across MuLVs (**Fig. 3.12 b**). Additionally, we also tested whether MuLVs derepression influenced the chromatin environment in neighbouring genes (*Tubb3* and *Def8* are shown as an example in **Fig. 3.12**). Interestingly, some of the MuLVs flanking genes were minimally covered by H3K9me3 in control and mutant cells; however we could appreciate a significant increase of the active H3K4me3 mark in *Setdb1*<sup>Mb1</sup> pro B cells also across these regions (**Fig. 3.12 b**).

All together these data demonstrated that SETDB1 binds to MuLV elements and that these elements are covered with H3K9me3 in control pro B cells. Nevertheless, it remains unclear whether MuLVs silencing is mediated via establishment of H3K9me3 or if this mark is just a consequence of SETDB1-mediated silencing. Notably, we observed that *Setdb1* depletion results in a decrease but not a complete loss of H3K9me3 across MuLV elements. Presence of

## Results (II)

H3K9me3, however, does not seem to be a hindrance for H3K4me3 establishment, which clearly occurs through all MuLV regions and; notably, also spread across their neighbouring genes such as *Tubb3* and *Def8* (**Fig. 3.12 b**).



**Figure 3.12 MuLVs partially lost the repressive H3K9me3 while gaining the H3K4me3 transcriptional mark in *Setdb1*<sup>Mb1</sup> pro B cells**

a) ChIP analysis performed on CD43<sup>+</sup> CD19<sup>+</sup> sorted pro B cells.  $\alpha$ -H3K9me3 and b)  $\alpha$ -H3K4me3 were used to identify loss or gain of these marks over MuLV elements and neighbouring regions.

### 3.1.5 MuLVs derepression results in activation of neighbouring genes

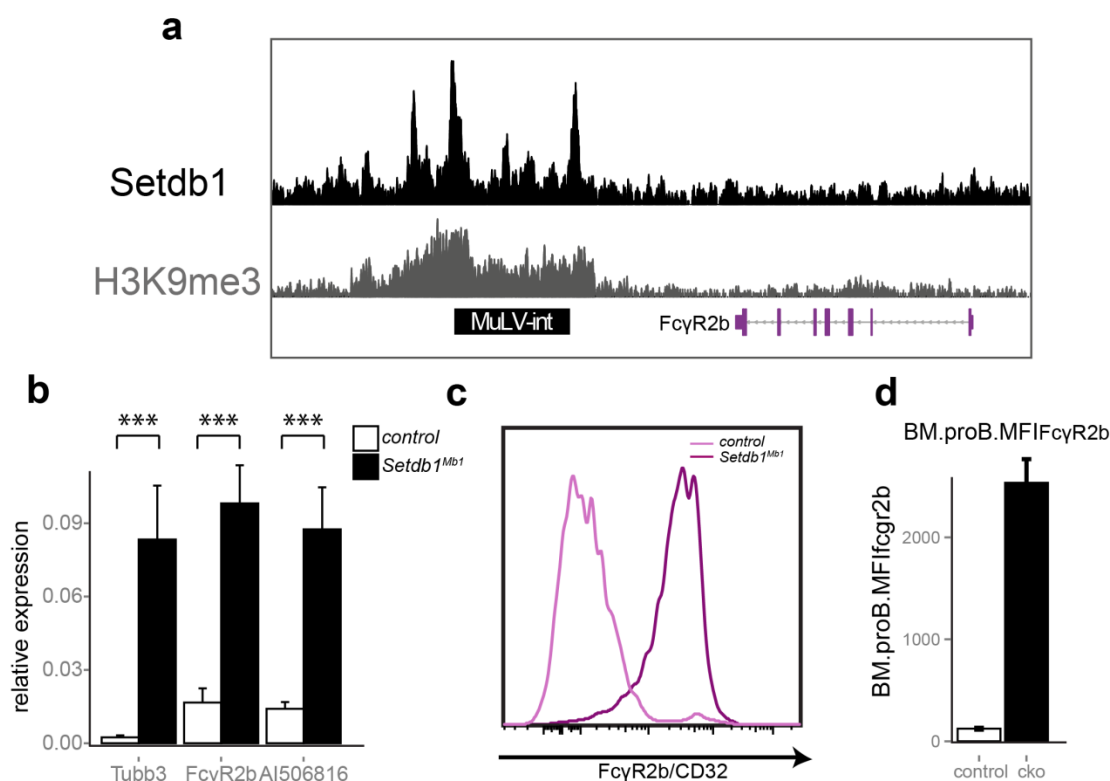
Because it is known that, beyond starting transcription of the proviral genome, LTRs promoters are able to enhance the transcription of genes which are located in their vicinity (Lu et al., 2014; van de Lagemaat et al., 2003); we verified whether transcription of MuLVs neighbouring genes was also affected in the absence of *Setdb1*. Plus, ChIP-qPCR analysis displayed a substantial gain of the active mark H3K4me3 over MuLV proximal genes, which might indicate the presence of a transcriptionally active environment also in the vicinity of MuLVs genomic regions (**Fig. 3.12 b**).

At first, we looked in our RNA-Seq profile at genes located at MuLVs 5' and 3' end. Interestingly, we found that all MuLVs flanking genes were highly transcribed, although neither SETDB1 nor H3K9me3 occupied their genomic sequence (**Fig. 3.13 a; 3.10 a**).



## Results (II)

Significantly, transactivation of MuLVs flanking genes was also confirmed via qRT-PCR in 6 biological replicates (**Fig. 3.13 b**).



**Figure 3.13 Transactivating functions of retrotransposons on neighbouring genes**

**a)** SETDB1 and H3K9me3 coverage over one representative MuLV region and its flanking gene. **b)** Overexpression of MuLV neighbouring genes in *Setdb1*<sup>Mbl</sup> sorted pro B cells (n=6). \*\*\*P < 0.001 (unpaired two-tailed Student's t-test). **c)** Representative FACS histograms showing FcγR2b protein detection on *Setdb1*<sup>Mbl</sup> pro B cell surface **d)** Bar graph displaying FcγR2b protein expression level calculated as average of the mean fluorescence intensity (n=6).

These findings indicate that upregulation of MuLVs neighbouring genes does not directly result from loss of SETDB1 repressive function, but rather comes from the enhancing role that LTR promoters exert in their proximity.

Taking advantage of the fact that one of MuLVs neighbouring genes was a cell surface marker, we tested by FACS analysis whether these transcripts were also translated. Consistently to what we observed for the viral protein env, also transactivated genes were translated into proteins, as it shown by FcγR2b overexpression onto the *Setdb1* deficient pro B cell surfaces (**Fig. 3.13 c and d**).

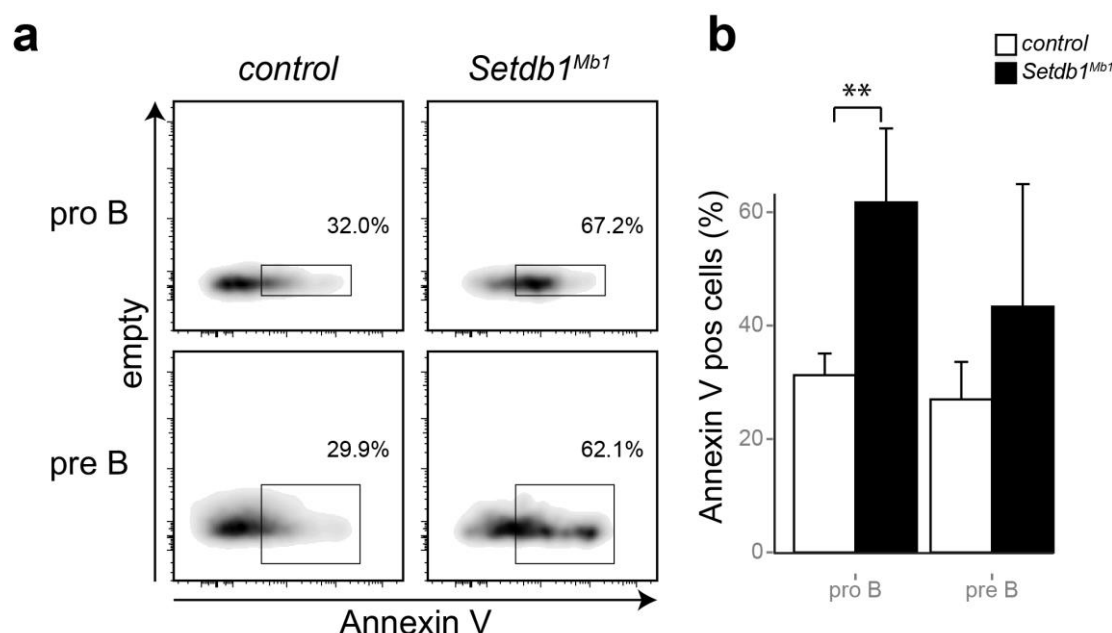
### 3.1.6 *Setdb1* loss in pro B cells induces apoptosis

It has already been demonstrated that accumulation of viral transcripts induces activation of the interferon pathway which triggers apoptosis. This happens because retroviral endogenous transcripts are somehow recognized by molecules that specifically detect viral DNA or RNA,

## Results (II)

inducing programmed cell death (Pichlmair and Reis e Sousa, 2007; Stetson and Medzhitov, 2006). Other evidences also demonstrated that expression of the viral envelope protein on the cell surface can induce programmed cell death via several mechanisms (Biard-Piechaczyk et al., 2000; Perfettini et al., 2005; Ullrich et al., 2000).

In order to understand whether ectopic expression of endogenous retroviruses induces apoptosis in pro B cells, we tested *Setdb1*<sup>Mb1</sup> pro B cells for apoptotic markers. Annexin V staining revealed that 30% of control pro B cells underwent apoptosis, reflecting the rate of aberrant IgH recombination which occurs at this stage of development. Interestingly, when we tested *Setdb1*<sup>Mb1</sup> pro B cells, we observed enhanced apoptosis: nearly the double compared to control cells (**Fig. 3.14 a and b**).



**Figure 3.14** *Setdb1*<sup>Mb1</sup> pro B cells showed enhanced apoptosis

**a)** Apoptosis detected via AnnV staining. **b)** Statistics from **a)** (n=6). \*\*P < 0.01 (unpaired two-tailed Student's t-test).

### 3.1.7 Overexpression of antiapoptotic *Bcl2* partially rescues B cell defects in *Setdb1*<sup>Mb1</sup> mice

Nearly all *Setdb1*<sup>Mb1</sup> pro B cells undergo programmed cell death, as demonstrated by the Annexin V staining (**Fig. 3.14 a and b**). To understand whether this phenotype could be rescued by overexpression of antiapoptotic molecules, we generated *Setdb1*<sup>Mb1</sup>; *VavBcl2* mice, which overexpress the antiapoptotic *Bcl2* throughout all the haematopoiesis. Remarkably, MACS sorted lineage negative *Setdb1*<sup>Mb1</sup>; *VavBcl2* progenitors readily formed pre B cell colonies in methylcellulose medium supplemented with IL-7, in a comparable fashion to

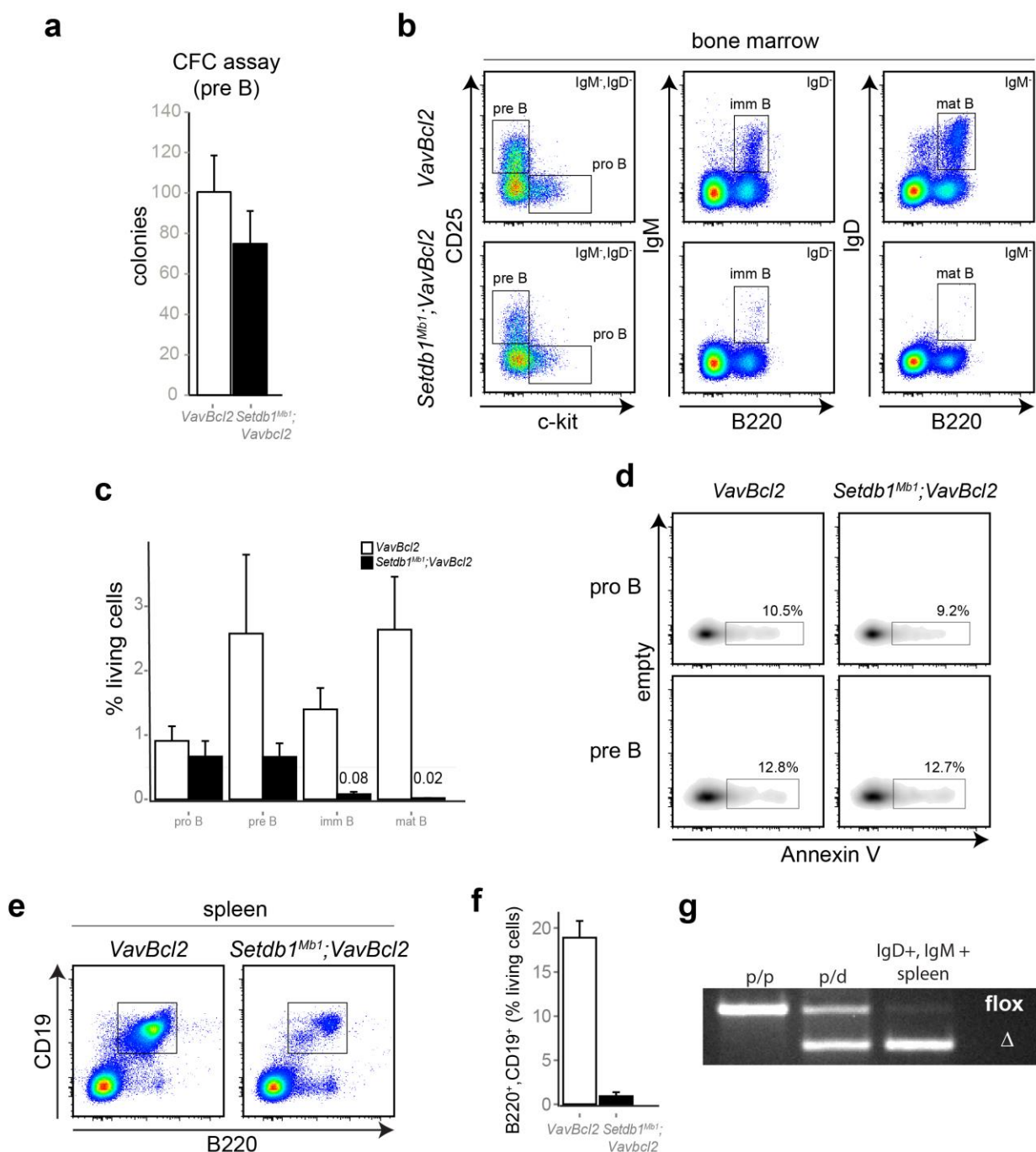


## Results (II)

*VavBcl2* control (**Fig. 3.15 a**). Although *Setdb1<sup>Mb1</sup>* B cell differentiation defects could be rescued *in vitro*, we also investigated whether *in vivo Bcl2* overexpression could bypass *Setdb1<sup>Mb1</sup>* pro B cell apoptosis. Therefore, we analysed B cell development in *Setdb1<sup>Mb1</sup>; VavBcl2* bone marrow. We performed immunophenotyping to specifically check all B cell developmental stages. As shown in **Fig. 3.15 b**, pro B cell numbers in *Setdb1<sup>Mb1</sup>; VavBcl2* mice were comparable to *VavBcl2* controls and, importantly, pro B cells underwent very little apoptosis, proving that *Bcl2* overexpression clearly rescues cell death at the pro B cell stage (**Fig. 3.15 d**).

Although *Bcl2* overexpression was obviously not enough to fully restore B cell development in *Setdb1<sup>Mb1</sup>* mice (**Fig. 3.15 c**), we observed that *Setdb1<sup>Mb1</sup>; VavBcl2* pre B cells increased 5 fold compared to *Setdb1<sup>Mb1</sup>* mice without *VavBcl2* transgene; demonstrating that in the absence of *Setdb1* pro B to pre B cell transition can be partially restored preventing apoptosis (**Fig. 3.15 b and c; Appendix Fig. 5.1**). Analysis of later stages of B cell development in *Setdb1<sup>Mb1</sup>; VavBcl2* mice, also showed an increase in immature B and mature B cells compared to *Setdb1<sup>Mb1</sup>* mice, which completely lacked mature B cells (**Fig. 3.15 b and c; Fig. 3.4 a and b**). Notably, these populations were able to migrate to the spleen (**Fig. 3.15 e and f**), demonstrating that even though B cell development has not completely recovered in *Setdb1<sup>Mb1</sup>; VavBcl2* mice, we can detect differentiated B cells also in the periphery. Importantly, we showed that these cells largely deleted *Setdb1*, indicating that splenic *Setdb1<sup>Mb1</sup>; VavBcl2* B cells were not deletion escapers (**Fig. 3.15 g**).

Altogether these data provide *in vitro* and *in vivo* evidences that *Bcl2* overexpression in *Setdb1<sup>Mb1</sup>* mice results in partial rescue of B cell development, indicating that many *Setdb1<sup>Mb1</sup>* B cells undergo apoptosis due to reasons which are not strictly linked to developmental defects. Developmental failures, in fact, cannot be rescued by just overexpressing antiapoptotic molecules, because cells would be kept anyway in a more immature state missing the tools to progress to the next developmental stage. Pro B to pre B cell transition in *Setdb1<sup>Mb1</sup>; VavBcl2* mice is, however, only moderately resolved; therefore we cannot exclude the possibility that *Setdb1* plays additional roles during early and/or later stages of B cell development.



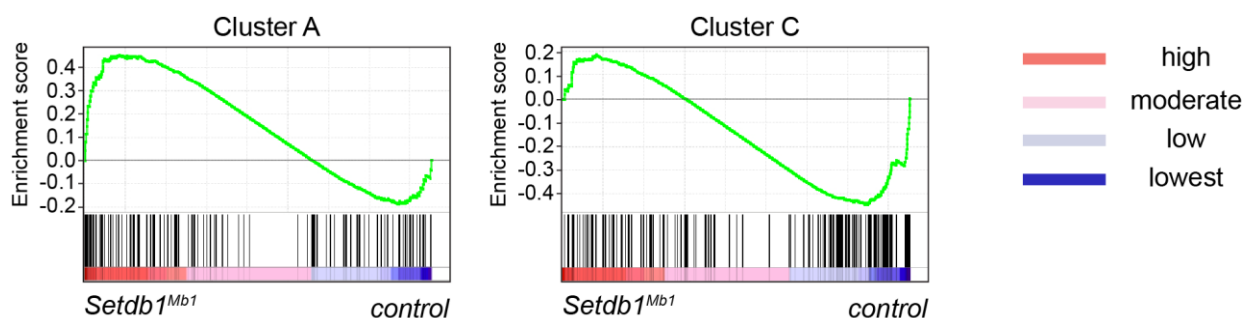
**Figure 3.15 Enforced *Bcl2* expression partially rescues B cell development in *Setdb1<sup>Mb1</sup>* pro B cells**  
**a)** pre B cell colony assay. *VavBcl2* and *Setdb1<sup>Mb1</sup>; VavBcl2* Lin<sup>-</sup> bone marrow cells were grown on methylcellulose medium containing IL-7 for 10 days. Colony number represents the average of 3 biological replicates. **b)** B cell developmental rescue in *Setdb1<sup>Mb1</sup>; VavBcl2* mice. FACS analysis of the main B cell populations in the bone marrow: pro B (CD19<sup>+</sup>, IgM<sup>-</sup>, IgD<sup>-</sup>, CD25<sup>-</sup>, c-kit<sup>+</sup>), pre B (CD19<sup>+</sup>, IgM<sup>-</sup>, IgD<sup>-</sup>, CD25<sup>+</sup>, c-kit<sup>-</sup>), immature B (B220<sup>+</sup>, IgD<sup>-</sup>, IgM<sup>+</sup>), mature B (B220<sup>+</sup>, IgD<sup>+</sup>, IgM<sup>+</sup>). **c)** Statistics from b) calculated as percentage of living cells from 5 biological replicates. **d)** AnnV staining to detect apoptosis in *VavBcl2* and *Setdb1<sup>Mb1</sup>; VavBcl2* pro B cells (n=7). **e)** B cell (CD19<sup>+</sup>, B220<sup>+</sup>) population in the spleen (n=4). **f)** Statistics from e) calculated as percentage of living cells. **g)** *Setdb1* deletion rate in IgM<sup>+</sup> IgD<sup>+</sup> sorted splenocytes. Genomic DNA purified from sorted IgD<sup>+</sup> IgM<sup>+</sup> splenic B cells was amplified by PCR to test the presence of the floxed and  $\Delta$  *Setdb1* alleles.

### 3.2 Discussion (II)

SETDB1 is an essential histone methyltransferase during development. *Setdb1* *ko* mice die at the peri-implantation stage and conditional deletion of the enzyme during neuronal and mesenchyme development severely impairs cell differentiation programs (Dodge et al., 2004). So far, all the efforts to reveal why SETDB1 mediated silencing is so essential during developmental transitions have not provided a clear answer. This is due to the main limitations which come along *in vivo* studies, as lack of biological material due to lethal effects of *Setdb1* deletion, together with obvious difficulties to monitor cell behaviour in living organisms.

The best known developmental context is the haematopoiesis. Over the past years a lot of progresses in dissecting all transitions which occur during hematopoietic differentiation have been made. In fact, gain and loss of function experiments revealed the essential role of many transcription factors at every stage of hematopoietic development, which allowed drawing the actual hierarchical model of haematopoiesis (**Fig. 1.4**). Additionally, the expression of stage specific surface markers, represent a big advantage to isolate definite hematopoietic subpopulations directly from the bone marrow. Here, we induced *Setdb1* conditional deletion during B cell differentiation using the stage specific *Mb1cre* recombinase. Remarkably, this resulted in a strong block during pro B to pre B cell transition. Virtually, *Setdb1*<sup>*Mb1*</sup> mice produced pro B cells comparably to control mice, while the pre B cell population was reduced by 50 fold and none of these cells made it to the next stages of B cell development. Such strong blockage is similarly observed in mice lacking B cell master regulators; therefore we first thought that loss of *Setdb1* resulted in inability to establish the B cell program. This assumption was also supported by the fact that SETDB1 has been described as a silencer of developmental gene in mES cells (Bilodeau et al., 2009), although it is not yet clear whether this is the main silencing activity exerted by the enzyme. To check how *Setdb1* depletion affects transcription, we performed RNA-Seq in sorted pro B cells. We found 260 deregulated genes which equally divided in downregulated and upregulated. Considering the postulated silencing function of SETDB1 this observation was really surprising, as we expected massive gene derepression. Even more surprising was the fact that all main factors which are known to be essential for B cell development, as *Pax5*, *Ebf1*, *E2a*, *Rag1/2* were transcriptionally unaltered, indicating that the main transcriptional networks were successfully settled in the absence of *Setdb1*. Although none of these master regulators was affected, GO analysis highlighted that many deregulated transcripts clustered with high score in the category of immune regulators. To better understand whether gene deregulation was the cause or the consequence of the phenotype, we used *in vitro* expanded pro B cells to perform ChIP-Seq

analysis with  $\alpha$ -SETDB1 and  $\alpha$ -H3K9me3 to find SETDB1 target genes. Peak calling analysis revealed that SETDB1 set on many promoters; however at nearly all these sites H3K9me3 was absent. Importantly, we observed that exclusive presence of SETDB1 over TSSs did not affect gene regulation, as demonstrated by the minimal transcriptional changes which occurred at these sites in *Setdb1*<sup>Mb1</sup> pro B cells (**Fig. 3.16**).



**Figure 3.16 Exclusive SETDB1 occupancy over TSSs does not affect gene regulation**

Gene set enrichment analysis performed on genes sitting within cluster A and cluster B genomic areas derived from the ChIP-Seq profile (see **Fig. 3.11 a**).

Then, we asked our profile which genomic regions were co-occupied by SETDB1 and H3K9me3. Repetitive elements and intergenic regions resulted to be the areas where SETDB1 and H3K9me3 showed the strongest correlation, possibly indicating that SETDB1 repressive function are exerted at these sites. To verify this, we checked in our RNA-Seq profile whether these regions were derepressed in *Setdb1*<sup>Mb1</sup> pro B cells. Importantly, we found that two classes of endogenous retroviruses, known as murine leukaemia viruses (MuLVs) and mouse VL30 retroelements (MMVL30), were significantly transcribed. Noteworthy, in mES cells loss of *Setdb1* resulted in specific upregulation of another class of element known as internal A-particles (IAPs) but not MMVL30 and MuLVs. The specific activation of certain classes of retroviruses in different *Setdb1* deficient cell types, clearly indicated that although SETDB1 covers these elements, their transcription is not simply a function of *Setdb1* depletion but also requires the presence of tissue specific transcription factors which are able to start their transcription. For example, it has been shown that in certain kind of lymphoid neoplasms the B cell factor E2A binds to MuLVs elements influencing their transcription (Lawrenz-Smith and Thomas, 1995). In this regard, further experiments could be performed to investigate which cell specific transcription factors are required to selectively activate transcription of distinct class of retrotransposons. Moreover, it would be interesting to verify whether ectopic expression of *E2a* in *Setdb1*ko ESCs or MEFs is able to induce MuLV transcription.

Another interesting aspect about transcription of endogenous elements regards the chromatin environment which needs to be established to allow their activation. In fact, we observed in

## Discussion (II)

---

our ChIP-qPCR experiment that upon *Setdb1* depletion H3K9me3 marks were only reduced, but not completely lost. This can be explained by the fact that this methyl mark is removed by passive demethylation, a process which requires cell cycle to occur. Because pro B cell population contains slow-cycling cells due to the VDJ recombination, it is reasonable to think that a complete loss of the H3K9me3 is not achievable due to the low number of cell divisions occurring at this stage. Interestingly, despite the presence of residual H3K9me3, the active H3K4me3 mark could be established, therefore complete removal of H3K9me3 is not required for H3K4me3 installation. This finding questions the role of chromatin modifications as sole regulators of transcription. In fact, it would be interesting to verify whether reintroducing a catalytically inactive form of SETDB1 inside *Setdb1*<sup>Mbl</sup> pro B cells would rescue the B cell developmental blockage.

Retroviral elements, together with other silenced genomic regions, were depicted in former times as non-coding or 'junk' DNA. With these terms biologists indicated genomic areas that could never be linked with any function in cell physiology. However, in the light of the most recent discoveries this concept needs to be reconsidered. Indeed, contemporary studies demonstrated that endogenous retroviral elements enhanced the expression of key developmental genes in embryonic stem cells (ESCs), suggesting a clear physiological role of these sequences in transcriptional regulation (Lu et al., 2014; Ryan, 2004). In our work, we also observed that derepression of endogenous retroviruses upon *Setdb1* loss resulted in transactivation of neighbouring genes. In fact, *Setdb1*<sup>Mbl</sup> pro B cells transcribed genomic areas which were located next to the retroelements; although in wild type these regions were not covered by SETDB1 or H3K9me3. Importantly, some of them showed opposite transcriptional orientation in respect to their associated repetitive elements, excluding the possibility that fusion proteins could originate from reading-through effects of the polymerase.

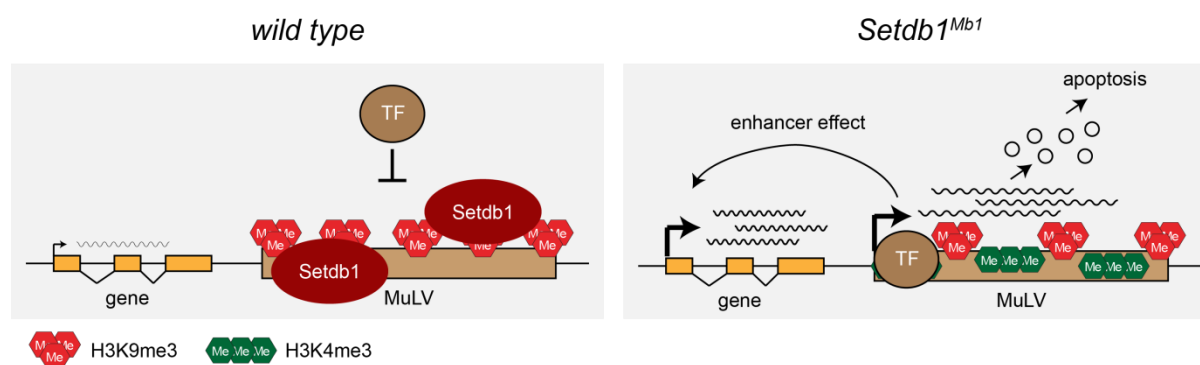
Ectopic expression of MuLVs has been associated with programmed cell death due to cytotoxicity induced by ER stress (Nanua and Yoshimura, 2004; Yoshimura et al., 2001; Zhao and Yoshimura, 2008). Additionally, expression of the env protein in peripheral human B cells has been found to interfere with cell proliferation (Jelicic et al., 2013). These evidences indicate that derepression of endogenous elements damages cell metabolism. Therefore we checked whether *Setdb1*<sup>Mbl</sup> pro B cells underwent apoptosis, independently from physiological cell death associated to the VDJ combinatorial events. *Setdb1* deficient pro B cells showed 2 fold more apoptosis compared to control cells, which exhibited a physiological 30%. Notably, if B cell apoptosis due to *Setdb1* depletion originated from developmental defects (such as impaired immunoglobulin rearrangement), it would not be possible to rescue differentiation via enforced expression of antiapoptotic due to the fact that B cell development impose

## Discussion (II)

obligated checkpoints which cannot be bypassed. To address this point, we generated *Setdb1<sup>Mb1</sup>; VavBcl2* mice, which overexpressed antiapoptotic *Bcl2* during haematopoiesis. These animals showed a partial but substantial rescue of B cell development, as they produced 5 fold more pre B cells compared to *Setdb1<sup>Mb1</sup>* mice. Importantly, *Setdb1<sup>Mb1</sup>; VavBcl2* pre B cells were also able to proceed further during B cell development and relocate in peripheral organs. Indeed, these cells populated the spleen and, notably, they were all *Setdb1* deficient. The partial rescue observed *in vivo* was a clear demonstration that part of *Setdb1<sup>Mb1</sup>* deficient pro B cells could bypass apoptosis upon *Bcl2* overexpression, continuing their path though B cell development. However, the fact that some cells still struggled to proceed to next developmental stages strongly suggests that of SETDB1 might have additional roles which cannot be resolved by *Bcl2* overexpression or that these cells follow a different pathway which leads to death.

Recently, it has been observed that mice lacking antibody producing B cells, such as *Rag* deficient mice, significantly transcribed retroviral elements in macrophages. In this study Young *et al.* showed that immune stimuli trigger expression of retroviral element which randomly reintegrated in the genome, generating instability and ultimately leukaemia. Interestingly, expression of these elements in mature differentiated cells as macrophage seemed not to affect cell viability, while in our study pro B cells were dramatically affected by the presence of retroviral transcripts. This might indicate that developing cells cannot tolerate ectopic expression of retroviruses, differently from mature cells.

In this regards, it would be intriguing to generate conditional knockout mice which delete *Setdb1* in later stages of B cell development to understand (1) whether this also induces derepression of repetitive elements, (2) if mature B cells can cope with their presence.



**Figure 3.17 Model**

In summary our data support the role of SETDB1 as main silencer of retroviral elements in the pro B cells genome. We observed that loss of *Setdb1* results in derepression of retroviruses

## Discussion (II)

---

without affecting the establishment of B cell transcriptional networks. Therefore, we assume that ectopic transcription of MuLVs and MMVL30 strongly interferes with B cell development inducing cellular stress which ultimately results in apoptosis (**Fig. 3.17**). Our results propose the first molecular link between *Setdb1* depletion and developmental defects, highlighting the importance of silencing endogenous retrovirus during B cell development.

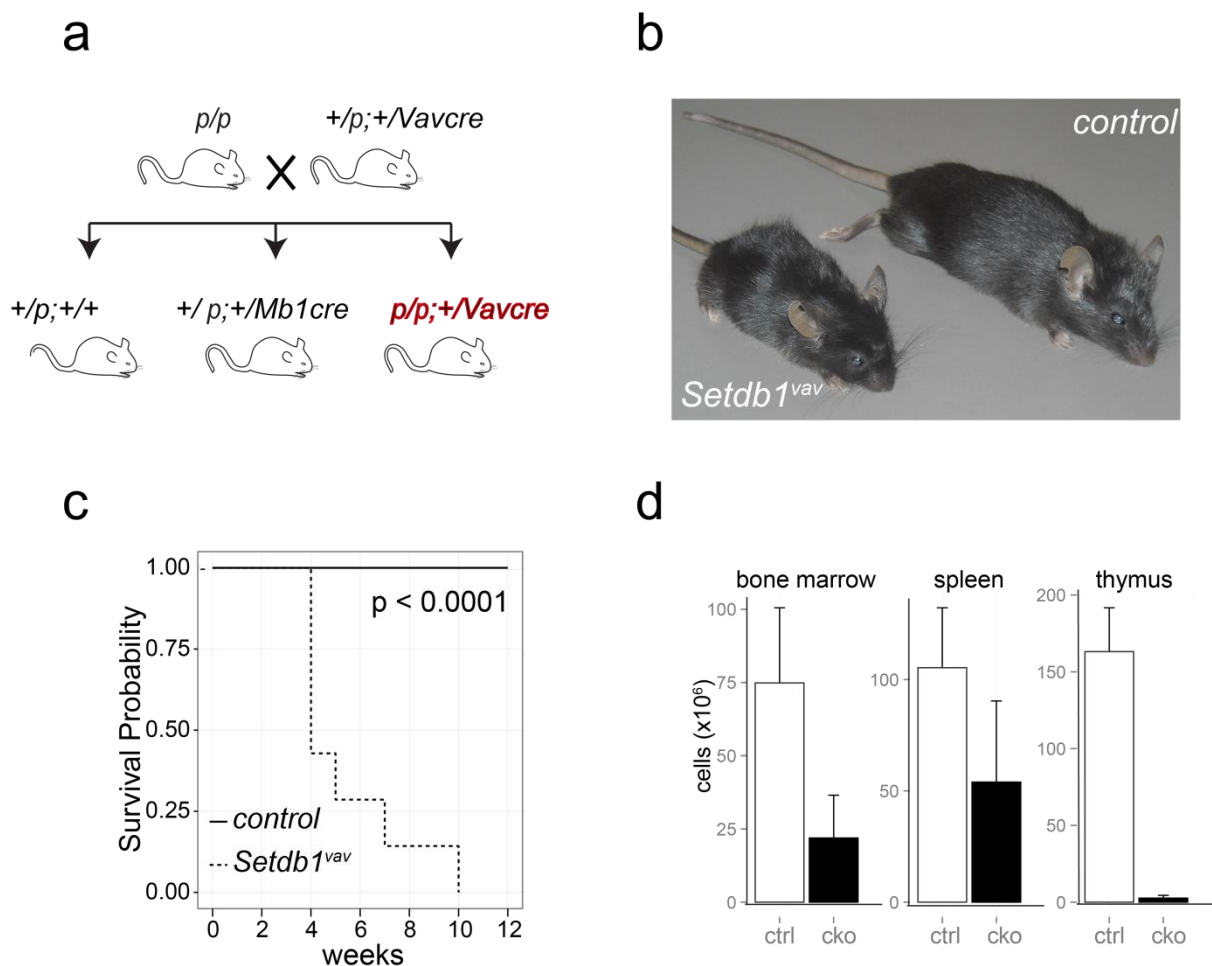


## 4 Results (III)

### 4.1 The role of the histone methyltransferase SETDB1 during haematopoiesis

#### 4.1.1 *Setdb1*<sup>Vav</sup> mice are underdeveloped and show impaired hematopoietic organs

To assess the role that SETDB1 plays during early stages of haematopoiesis, we generated conditional knockout mice (*Setdb1*<sup>Vav</sup>). These animals were obtained by crossing *Setdb1*<sup>p/p</sup> mice, which carried two *Setdb1* floxed alleles with *VavCre* mice, accomplishing tissue specific *Setdb1* depletion in all hematopoietic cells (**Fig. 4.1 a**).



**Figure 4.1** *Setdb1*<sup>Vav</sup> mice were underdeveloped and show hematopoietic defects

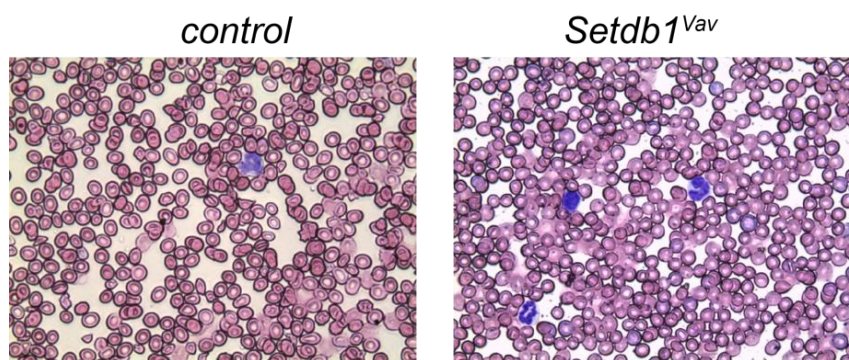
**a)** Breeding scheme adopted to obtain control and *Setdb1*<sup>Mb1</sup> cko mice. **b)** Underdeveloped *Setdb1*<sup>Vav</sup> mice. All analysed animals weighted 2-3 folds less than controls and died within 4<sup>th</sup> week of age **c).** **d)** Hematopoietic organs cellularity calculated from at least 6 biological replicates.

*Setdb1*<sup>Vav</sup> 4 weeks old mice were 3 fold smaller compared to littermate controls (**Fig. 4.1 b**) and they showed clear developmental defects. In facts, in addition to reduced body size,



## Results (III)

mutant mice exhibited streaked tails, showed a sparse hair coat, lack of down teeth and twisted upper teeth. To make up for the lack of down teeth, mice were provided with wet diet soon after weaning or they were kept in the mating cages for a longer time. Also, we verified that weaned mice could chew and eat properly, to avoid further complication which might have interfered with the already existing developmental defects. Unfortunately, even though *Setdb1*<sup>Vav</sup> mice could be regularly fed and behaved normally, they died within the first 4-5 weeks after birth (**Fig. 4.1 c**). At first, we associated early postnatal lethality to lack of erythrocyte. However, this hypothesis was excluded when we analysed *Setdb1*<sup>Vav</sup> blood smear and observed that red blood cells were regularly produced and that their morphology was unchanged compared to control (**Fig. 4.2**).



**Figure 4.2 *Setdb1*<sup>Vav</sup> erythrocytes are morphologically unaltered**  
Blood smears from control and *Setdb1*<sup>Vav</sup> mice stained May-Grünwald GIEMSA.

Due to short survival time, we carefully monitored all littermates in this colony to identify mutant mice even before weaning. As soon as *Setdb1*<sup>Vav</sup> mice could be identified and weaned, they were immediately scarified and prepared for analysis of all hematopoietic organs.

The strong developmental defects shown by *Setdb1*<sup>Vav</sup> mice already heralded the possibility of a severe impairment of the hematopoietic system. Indeed, when we started analysing the hematopoietic organs, we first noticed that *Setdb1*<sup>Vav</sup> mice exhibited a rudimental thymus which was hardly visible with naked eyes. Similar size reduction was observed in other lymphoid organs, like mesenteric lymph nodes and Peyer's patches, which could be hardly identified in the abdomen of these animals. Surprisingly, *Setdb1*<sup>Vav</sup> spleens were comparable in weight to control spleens and they were enlarged 2-3 folds in proportion to body size. Bone marrow cellularity in *Setdb1*<sup>Vav</sup> mice was reduced by 3 fold (**Fig. 4.1 d**). Yet, this could just reflect the smaller size of the mice.

Altogether these developmental defects showed that tissue specific depletion of SETDB1 using the *VavCre* recombinase, resulted in severe underdevelopment and impairment of the hematopoietic compartment.

#### 4.1.2 Lack of *Setdb1* during hematopoietic development results in complete loss of lymphocytes and expansion of myeloid-erythroid lineage in the periphery

To further understand which steps of hematopoietic development broke down in the absence of *Setdb1*, we started analysing all terminally differentiated cell of the immune system.

Thymus is the central hematopoietic organs where T cells develop before migrating to the periphery. Young mice count approximately two million thymocytes, most of which are developing T cells. *Setdb1<sup>Vav</sup>* thymus size was reduced 100 folds compared to controls (**Fig. 4.1 d**), indicating that this organ could not really sustain proper T cell development (**Fig. 4.3 b**). On the other hand, spleen size was unaffected; therefore we asked which kind of cells populated *Setdb1<sup>Vav</sup>* spleens. Consistently with what we observed in the thymus, we could detect no splenic T cells, demonstrating that the few T cells produced are not able to reach peripheral organs (data not shown). Next, to test if also B cells were affected in the lymphoid compartment, we stained *Setdb1<sup>Vav</sup>* splenocytes with B cells markers. Notably, CD19+, B220+ cells were completely missing in *Setdb1<sup>Vav</sup>* spleens, showing that this organ was totally deprived of lymphoid cells (**Fig. 4.3 a bottom panel**). Excluding the presence of lymphoid cells in *Setdb1<sup>Vav</sup>* spleen, we checked for myeloid-erythroid populations, the other cell types that normally resides into the spleen. *Setdb1<sup>Vav</sup>* Gr+ Mac+ cells, which represent the myeloid fraction, were overrepresented and the macrophage (Mac+, Gr-) population was increased more than 3 fold (**Fig. 4.3 c bottom panel**). Additionally, we tested erythroid cells, which can be distinguished according to the developmental stage they are going through. Although we detected mild differences in each developmental steps, terminally differentiated erythrocytes were normally produced and looked morphologically unaltered (**Fig. 4.3 d**), as we could also appreciate in the GIEMSA staining performed on blood smears from *Setdb1<sup>Vav</sup>* mice (**Fig. 4.2**). In summary, we conclude that *Setdb1<sup>Vav</sup>* mice completely lack terminally differentiated B and T cells but overexpand myeloid-erythroid populations in the periphery.

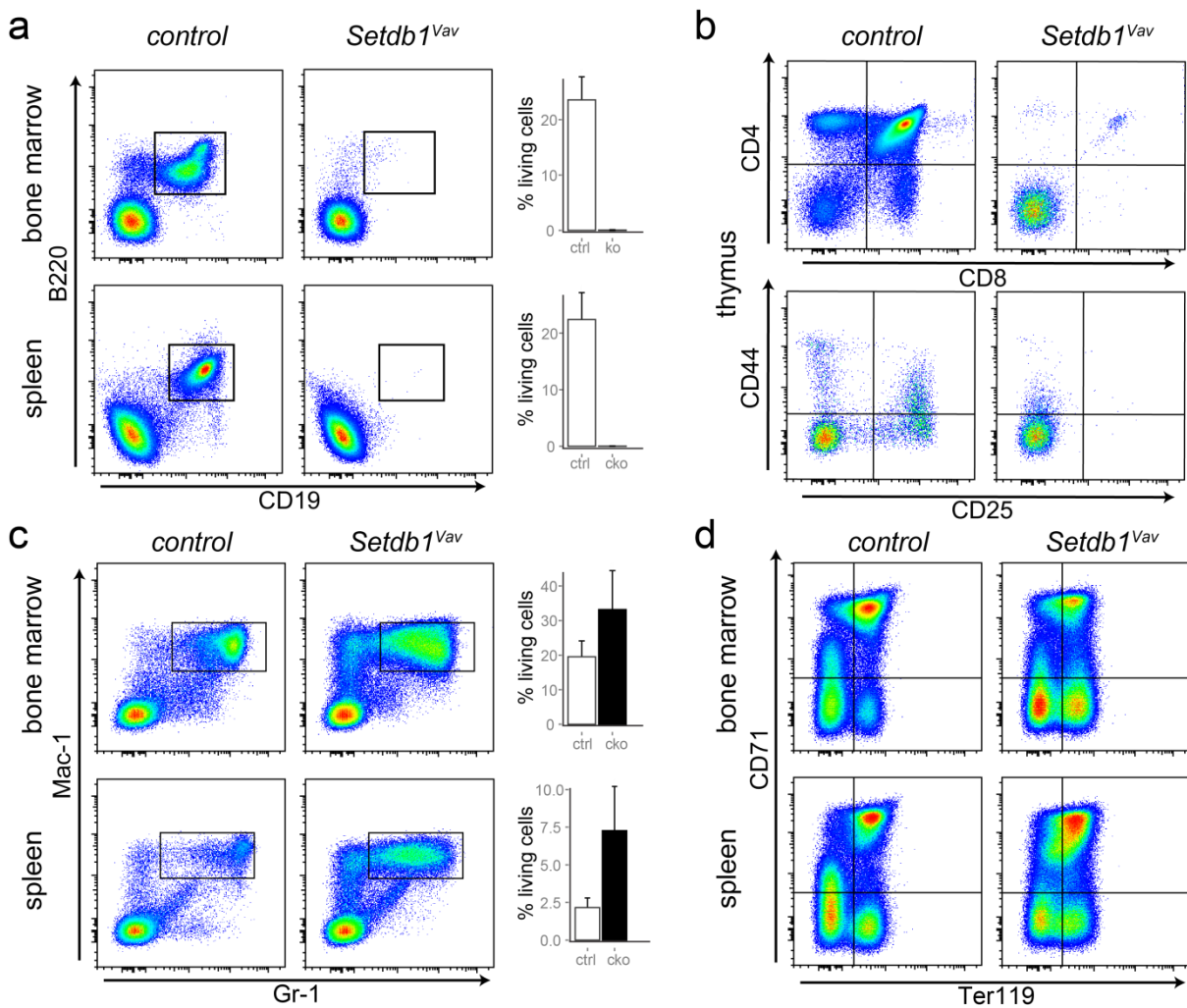
#### 4.1.3 *Setdb1<sup>Vav</sup>* bone marrow completely abolishes B cell development in favour of myeloid-erythroid expansion

Because we observed a total loss of lymphoid cells and a gain of myeloid population in the periphery, we also analysed how the myeloid and lymphoid compartment behaved in the bone marrow, headquarter of the haematopoiesis. Analysis of *Setdb1<sup>Vav</sup>* B cell population in the

## Results (III)

bone marrow confirmed what we have already observed in the spleen. In fact, also here CD19<sup>+</sup>, B220<sup>+</sup> cells were missing (**Fig. 4.3 a top panel**). Because of complete lack of B cell markers, we could not detect any population in all B cell differentiation steps, confirming that loss of *Setdb1* in early haematopoiesis completely abrogates B cells development. Because we observed an increase in Gr-1<sup>+</sup> Mac-1<sup>+</sup> cells and erythroblasts in the periphery, we also tested these populations in the bone marrow, which is the main myeloid organ.

Coherently to what we found in the spleen, myeloid cells and erythroblasts were overrepresented. In particular, we also noticed the presence of an unusually abundant Mac-1 only positive cell population which is produced in very low amount in the bone marrow (**Fig. 4.3 c bottom panel**).



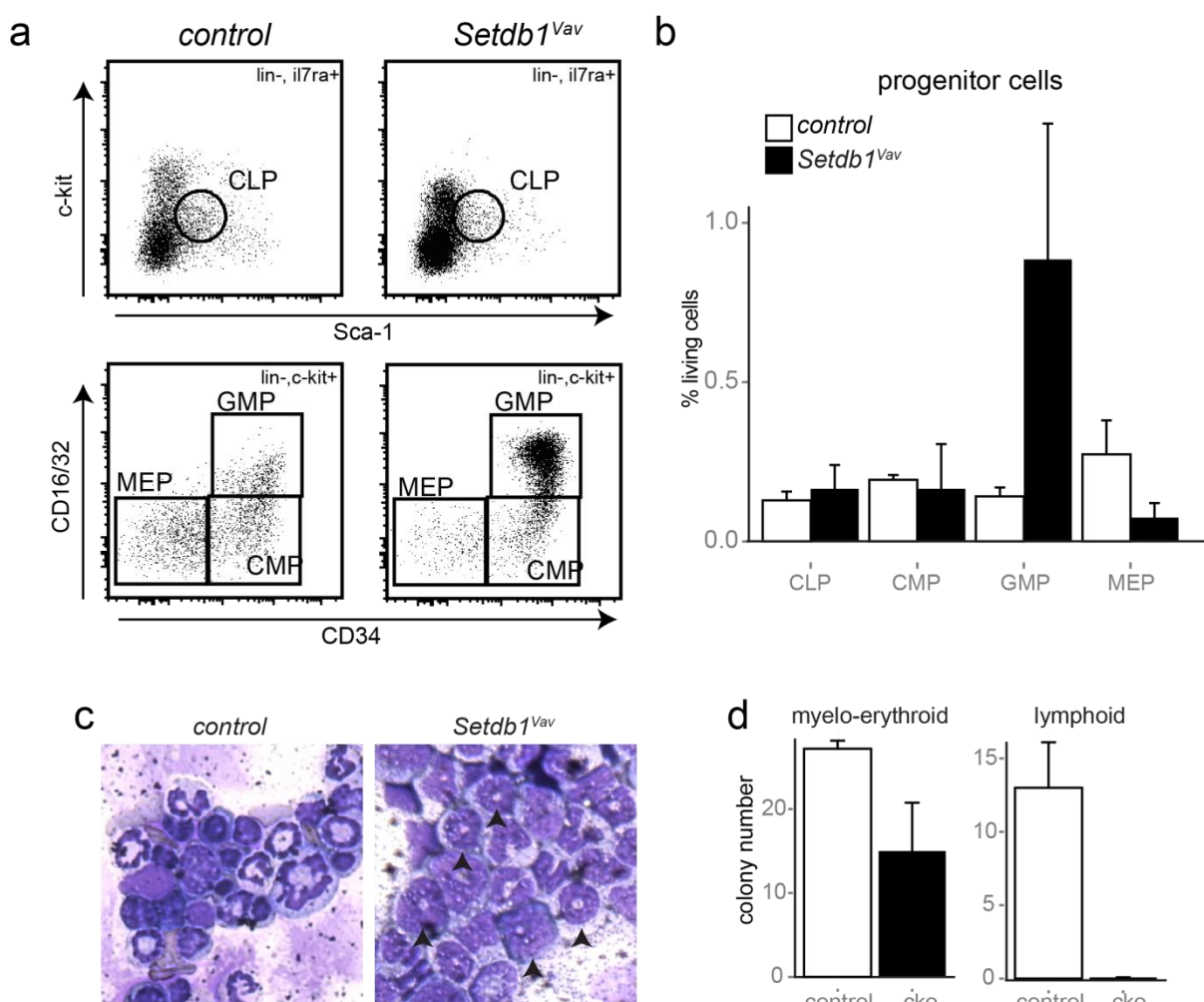
**Figure 4.3 Hematopoietic development is severely compromised in *Setdb1*<sup>Vav</sup> mice**

**a-d**) Flow cytometry analysis of bone marrow and splenic B cells (B220<sup>+</sup>, CD19<sup>+</sup>) and bargraph depicting average and standard deviation of 8 biological replicates **a**); main T cell populations in the thymus **b**); myeloid (Gr<sup>+</sup> Mac<sup>+</sup>) cell population together with bargraph depicting average and standard deviation of 8 biological replicates **c**) and erythroid cells **d**) in bone marrow and spleen.

#### 4.1.4 Common lymphoid progenitors (CLPs) are not altered in *Setdb1*<sup>Vav</sup> mice while granulocyte macrophage progenitors (GMPs) are overrepresented

In order to understand whether B cell failure and overexpansion of myeloid population in *Setdb1*<sup>Vav</sup> mice was due to aberrant production of hematopoietic progenitors, we performed FACS analysis to test lymphoid and myeloid precursors.

Common lymphoid progenitors (CLPs) give rise to B and T cells and some natural killer cells (NKs). Aberrant production of these cells would explain absence of lymphoid cells in *Setdb1*<sup>Vav</sup> mice. Surprisingly, *Setdb1*<sup>Vav</sup> CLPs were produced in a similar fashion to control mice (**Fig. 4.4 a top panel and b**), suggesting that this population was virtually present, but not able to further differentiate.



**Figure 4.4 Lymphoid progenitors are unaltered in *Setdb1*<sup>Vav</sup> bone marrow while myeloid progenitors show a bias towards GMPs differentiation**

**a)** Flow cytometry analysis of hematopoietic multipotent progenitors. Top panel shows CLPs (Lin<sup>-</sup>, IL7Rα<sup>+</sup>, Sca-1 low, c-kit low). Bottom panel depicts myeloid-erythroid progenitors: CMPs (Lin<sup>-</sup>, IL7Rα<sup>-</sup>, c-kit high, Sca-1<sup>+</sup>; CD34<sup>+</sup> and CD16/32<sup>-</sup>); GMPs (Lin<sup>-</sup>, IL7Rα<sup>-</sup>, c-kit high, Sca-1<sup>+</sup>, CD34<sup>+</sup> and CD16/32<sup>+</sup>); MEPs (Lin<sup>-</sup>, IL7Rα<sup>-</sup>, c-kit high, Sca-1<sup>-</sup>; CD34<sup>-</sup> and CD16/32<sup>-</sup>). **b)** Bargraphs depicting average and standard deviation of 6 biological replicates. **c)** Cytospin preparation of bone marrow cells stained with May-Grünwald Giemsa. **d)** Myeloid and lymphoid cell colony assay. Control and

## Results (III)

*Setdb1*<sup>Vav</sup> bone marrow progenitors were grown on methylcellulose medium enriched either with Stem Cell Factor IL-3 and IL-6 and Erythropoietin or IL-7 to promote myeloid-erythroid and lymphoid differentiation, respectively. Colony number represents the average of 3 biological replicates.

The overrepresentation of differentiated Gr<sup>+</sup> Mac<sup>+</sup> cells in the bone marrow and in the spleen for *Setdb1*<sup>Vav</sup> mice led us to investigate also myeloid progenitors. Myeloid precursors start to differentiate from the less mature common myeloid progenitors (CMPs), which then give rise to either megakaryocytes-erythrocytes progenitors (MEPs) or granulocyte macrophage progenitors (GMPs).

Percentages of *Setdb1*<sup>Vav</sup> CMPs were normal; however we appreciated a significant loss of MEPs and a substantial gain in GMPs (**Fig. 4.4 a bottom panel and b**). Although this clear bias towards GMPs production would nicely explain the overrepresentation of mature Gr<sup>+</sup> Mac<sup>+</sup> cells, loss of MEPs conflicts with the existence of normally shaped erythrocytes and normal amounts of erythroid precursors.

According to these data, we conclude that *Setdb1* depletion blocks lymphoid differentiation at the CLPs stage and generates a strong bias towards GMPs production.

### 4.1.5 HSCs compartment collapses 4 weeks after birth in *Setdb1*<sup>Vav</sup> mice

Hematopoietic stem cells are the starting point of all hematopoietic development, as they can initiate the specification of lymphoid and myeloid-erythroid lineage. To complete our view on hematopoietic development in *Setdb1*<sup>Vav</sup> mice we checked whether HSC population (Lin<sup>-</sup> Sca<sup>+</sup> c-kit<sup>+</sup> or LSK) was affected. Surprisingly, HSCs were not detectable in 4 weeks old bone marrow, indicating that 4 weeks old *Setdb1*<sup>Vav</sup> mice completely miss the cells which give rise to the entire hematopoietic compartment.

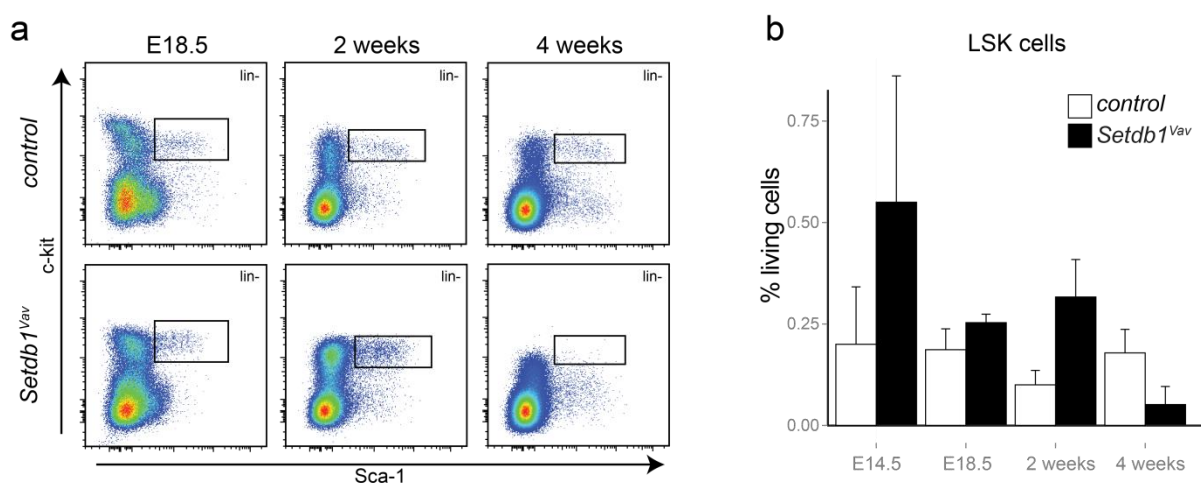
Because *Setdb1*<sup>Vav</sup> mice produced significant amounts of myeloid cells, we were wondering whether they differentiated from LSKs previously generated. To verify this hypothesis, we tested foetal liver from E18.5 embryos and bone marrow from 2-3 weeks old mice for the presence of hematopoietic stem cells. Interestingly, LSK population was increased 2 fold in both *Setdb1*<sup>Vav</sup> foetal liver cells and 2-3 weeks old bone marrow (**Fig. 4.5 a and b**). This finding demonstrate that *Setdb1*<sup>Vav</sup> mice produce enough LSKs in the foetal liver and during the first 3 weeks of life, while after the 4<sup>th</sup> week of age they manifest a full collapse of the HSCs compartment.

Importantly, comparably to what we observed in 4 weeks old *Setdb1*<sup>Vav</sup> mice, we could confirm lack of lymphoid cells and overexpansion of mature and immature myeloid cells in *Setdb1*<sup>Vav</sup> foetal liver and 2-3 weeks old bone marrow (data not shown); although at this stage the HSCs compartment was virtually not affected.



## Results (III)

These data suggest that loss of *Setdb1* might either induce HSCs exhaustion or, alternatively, impaired maintenance of the HSCs.



**Figure 4.5 HSCs compartment collapses in *Setdb1*<sup>Vav</sup> mice 4 weeks after birth**

**a)** Flow cytometry analysis of hematopoietic stem cell population in E18.5 foetal liver (left), 2 weeks (middle) and 4 weeks old bone marrow (right). Hematopoietic stem cells were identified by FACS analysis as Lin<sup>-</sup>, Sca<sup>+</sup>, c-kit<sup>+</sup> (LSK). **b)** Bargraphs depicting average and standard deviation of 6 biological replicates.

## 4.2 Discussion (III)

In the previous discussion part we have already spoken about the importance of the histone methyltransferase *Setdb1* during development. Moreover we described how loss of *Setdb1* interferes during B cell differentiation due to aberrant expression of retrotransposable elements. These data strongly supported the idea that developmental defects observed in *Setdb1*<sup>Mb1</sup> mice are linked to retrotransposons induced toxicity. If this was always the case, broad deletion of *Setdb1* throughout all hematopoietic development would have completely abolished the hematopoietic system due to global apoptosis. To address this point and to generally investigated *Setdb1* functions during all haematopoiesis, we generated *Setdb1*<sup>Vav</sup> mice. These animals induced excision of *Setdb1* floxed alleles already during foetal haematopoiesis when the recombinase starts to be expressed under the control of *Vav* promoter. *Setdb1*<sup>Vav</sup> mutant mice were not produced according to Mendelian ratio, but they were underrepresented compared to the other genotypes. This could be either due to (1) low penetrance of the *Vavcre* recombinase in our mouse colony; (2) to the fact that *Vav* promoter is also active in trophoblast cells, therefore interfering with embryo development (3) to high level of postnatal death. Although *Setdb1*<sup>Vav</sup> mice were produced with low frequencies, we managed to accumulate some interesting data about their phenotype.

*Setdb1*<sup>Vav</sup> mice were severely underdeveloped as they weighed 3 folds less than their littermate controls. Curiously, before weaning they had no problem in ambulation and ate normally, although their upper teeth were twisted and the lower ones were missing. After weaning, these mice needed a daily checking since they exhibited clear signs of weakness. For these reasons all mice used for analysis were taken before or soon after weaning to avoid that the hematopoietic phenotype was affected by other developmental problems. Noteworthy, *Setdb1*<sup>Vav</sup> mice died within 5 weeks after birth, as a result of a strong developmental phenotype (**Fig. 4.1 a**). If early death was the results of hematopoietic defects or the consequence of expression leakage of the recombinase in other tissue, it still has to be determined.

The essential role of *Setdb1* during haematopoiesis became immediately evident when we examined at the main hematopoietic organs. In fact, curiously, these mice completely lacked the thymus while spleen size was comparable to the controls. Bone marrow cellularity was also significantly decreased; however this effect was imputable to the fact that mice were smaller.

These macroscopic evidences were already suggestive of the pivotal role that *Setdb1* exerts during hematopoietic development.

## Discussion (III)

---

To better understand which subpopulations were affected in the absence of *Setdb1* we analysed all main hematopoietic lineages by FACS analysis. Remarkably, *Setdb1*<sup>Vav</sup> mice completely abrogated production of lymphoid cells, as demonstrated by total loss of T cells in thymus and missing B lymphocytes from bone marrow and spleen. Differently from the lymphoid lineage, the myeloid counterpart was overexpanded. In fact, either the bone marrow or the spleen were fully populated by myeloid-erythroid cells, indicating that *Setdb1* depletion induce a clear bias towards myeloid differentiation at the expenses of the lymphoid lineage. Notably, these observations led us to think that if loss of *Setdb1* induced toxicity due to aberrant expression of ERVs, all hematopoietic lineages should have been affected. However, this was not the case as myeloid-erythroid cells could tolerate *Setdb1* depletion. Moreover, GMPs overexpansion also suggested that *Setdb1* deficiency might push myeloid progenitor differentiation towards this direction.

Although *Setdb1*<sup>Vav</sup> haematopoiesis have pretty unique features, several mouse models which showed impaired lymphopoiesis, have a bias towards the myeloid lineage. For example, mice carrying *Dnmt1* hypomorphic alleles also show myeloid-erythroid restriction and aberrant lymphopoiesis (Broske et al., 2009). Also, absence of *Mef2c* hijacked haematopoiesis to the myeloid fate, severely impairing development of B, T and NK cells. In both papers, the molecular reasons which caused the phenotype were both found in the aberrant expression of master regulators which drive hematopoietic lineage decision. Indeed, *Mef2c* is the upstream transcription factor which regulates the myeloid factor PU.1 while DNA hypomethylation leads to myeloid-erythroid genes derepression in HSCs (Stehling-Sun et al., 2009). Unfortunately, we are not yet aware about molecular details which could explain the myeloid-erythroid expansion in *Setdb1*<sup>Vav</sup> mice; however ongoing experiments will hopefully clarify soon why *Setdb1* is so essential during haematopoiesis.

Because defects in mature blood cells might results from aberrant differentiation of progenitor cells, we started tracking back all hematopoietic development to check earlier developmental stages. Because *Setdb1*<sup>Vav</sup> mice did not generate any lymphoid cells we stained their direct precursor, the common lymphoid progenitors (CLPs). Surprisingly, these cells were regularly produced in *Setdb1*<sup>Vav</sup> bone marrow, indicating that at this stage *Setdb1* depletion abolished the establishment of B and T cell program.

Next, we also analysed myeloid progenitor population in *Setdb1*<sup>Vav</sup> bone marrow. Interestingly, common myeloid progenitors (CMPs), which can differentiate GMPs and MEPs, exhibited a clear preference towards GMPs differentiation, revealing that there is an evident trend to produce more Gr<sup>+</sup> Mac<sup>+</sup> cells already at the progenitor stage (**Fig. 4.4 a bottom panel**).

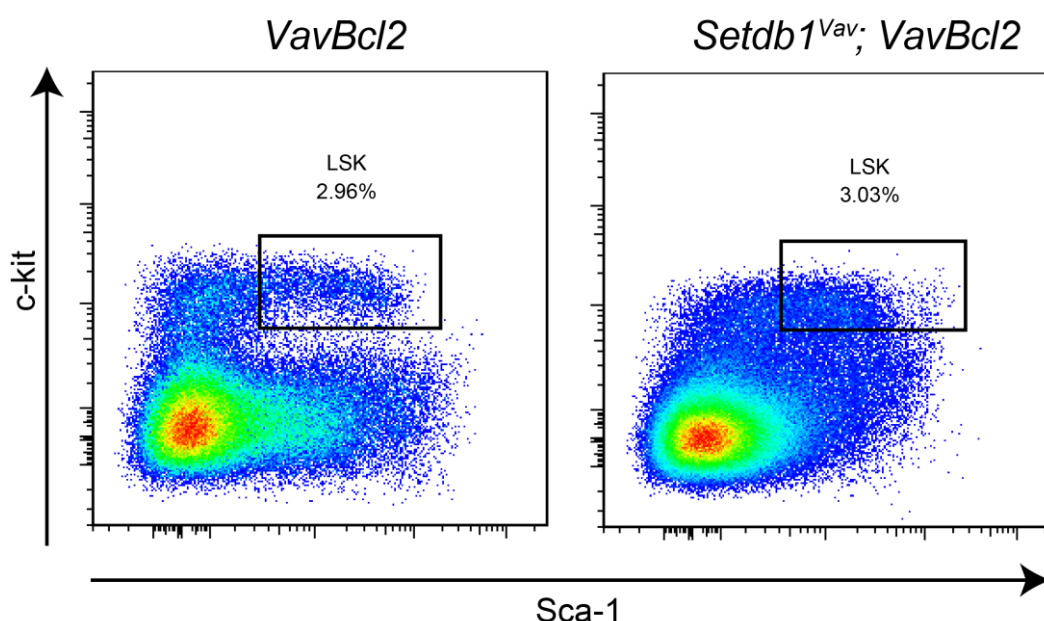


## Discussion (III)

Next, we verified whether LSK hematopoietic stem cells population was affected. Remarkably, we found that already after weaning these cells were severely underrepresented within the lineage negative progenitors and that c-KIT expression was markedly decreased (**Fig. 4.5**). Disappearance of LSK indicates an evident collapse of the haematopoietic system as from this population all blood cells originate; however we still do not know whether this phenotype can be lethal for the mouse or if additional developmental defects impaired LSKs maintenance in the hematopoietic stem cell niches. Because *Setdb1*<sup>Vav</sup> embryos or newly born *Setdb1*<sup>Vav</sup> mice did not show detectable developmental defects, we also analysed haematopoiesis in younger mice and in the foetus. This analysis was also performed because we could not explain the presence of mature hematopoietic cells in 4 weeks old mice while LSKs were missing. In fact, since all hematopoietic system originate from these cells their complete absence in younger mice or in the foetus would have resulted in global hematopoietic failure. As shown in **Fig. 4.5**, LSKs were regularly produced in the foetal liver or in 2-3 weeks old bone marrow, indicating that somehow *Setdb1* deficient LSKs can held hematopoietic development till the 4<sup>th</sup> week of age. Importantly, also CLPs, CMPs, MEPs and GMPs were checked in younger mice and all phenotype observed in 4 weeks old mice were confirmed as any B or T cell population was detectable and mature and progenitor myeloid cells were overrepresented (data not shown). Preliminary analyses to prove the cell intrinsic nature of the phenotype have been done *in vitro* using MethoCult 3630 and 3434 which promote lymphoid and myeloid differentiation, respectively. *Setdb1*<sup>Vav</sup> lineage negative progenitor cells failed to produce pre B cell colonies while they managed to produce some myeloid colonies (**Fig. 4.4 d**). This experiment suggested that the inability to generate lymphoid cells in the absence of *Setdb1* might be due to cell-intrinsic defects, however the myeloid overexpansion which characterize *Setdb1*<sup>Vav</sup> mice was not reproducible *in vitro*, since low amount of myeloid colonies were detectable. This effect could be attributed to the fact that we used progenitors from 4 weeks old mice; therefore LSKs function could have been already compromised. Colony formation assay for both myeloid and lymphoid potential will be repeated using Lin<sup>-</sup> progenitors derived from younger bone marrows where LSK population is still present. More importantly, we will also perform bone marrow transplantation experiments to test *in vivo* whether SETDB1 is intrinsically required for hematopoietic development. Unfortunately, the implementation of this experiment it will be particularly challenging since the number of CD4 CD8 depleted bone marrow cells that we can get from young *Setdb1*<sup>Vav</sup> mice is very limited. Alternatively, we could also inject mice with 5-FU to get progenitors out of *Setdb1*<sup>Vav</sup> bone marrows. This approach enriches for multipotent cells and gives the advantage to use lower amount cells for reconstitution. However, injection of *Setdb1*<sup>Vav</sup> mice,

## Discussion (III)

even at young age might be really harmful considering the strong developmental defects that these animals exhibit. Nevertheless, we hope to perform bone marrow transplantation as soon as enough biological material will be available. Meanwhile, we crossed *Setdb1*<sup>Vav</sup> mice with *VavBcl2* transgenic mice with the intent to rescue eventual apoptosis occurring in upon *Setdb1* depletion. Thus far, we managed to analyse only one *Setdb1*<sup>Vav</sup>; *VavBcl2* mouse. Interestingly, this animal clearly recovered the developmental weaknesses characteristic of *Setdb1*<sup>Vav</sup> mice, indicating that apoptotic rescue can improve *Setdb1*<sup>Vav</sup> mouse fitness. Additionally, we were able to age this mouse for 3 months and, during this period, no evident signs of weakness were detectable. To verify whether we could confirm the *Setdb1*<sup>Vav</sup> hematopoietic phenotypes we also analysed *Setdb1*<sup>Vav</sup>; *VavBcl2* by FACS. Remarkably, all defects detected in *Setdb1*<sup>Vav</sup> mice were reproduced in *Setdb1*<sup>Vav</sup>; *VavBcl2* animals. Indeed, lymphoid lineage differentiation was still not occurring and myeloid cells were clearly overexpanded. However; differently from *Setdb1*<sup>Vav</sup> mice, *Setdb1*<sup>Vav</sup>; *VavBcl2* transgenic LSKs were still present in 3 months old animals although, as shown in the FACS plot in **Fig. 4.6**, the marker distribution on the cell surface was clearly perturbed.



**Figure 4.6 LSK population from 3 months old *Setdb1*<sup>Vav</sup>; *VavBcl2* mice**  
Hematopoietic stem cells were identified by FACS analysis as Lin<sup>-</sup>, Sca<sup>+</sup>, c-kit<sup>+</sup> (LSK).

Significantly, more *Setdb1*<sup>Vav</sup>; *VavBcl2* mice have to be produced to confirm our preliminary data; yet these first observations encouraged us to believe that *Setdb1* is strongly required to balance lineage decision during hematopoietic development.

All data discussed so far are purely descriptive and all the assumptions we can make about the molecular mechanism which hides behind these phenotypes, remain merely speculative.

## Discussion (III)

---

Therefore, to elucidate the origin of all these defects, we planned to perform RNA-Seq analysis on sorted hematopoietic progenitors to check transcriptional changes occurring upon *Setdb1* depletion. Moreover, we could take advantage of the increased viability of *Setdb1*<sup>Vav</sup>; *VavBcl2* mice to perform aging experiment which will allow observing possible tumour onset due to persistent production of myeloid cells.

## 5 Appendix

### 5.1 Appendix (Results I)

**Table 5.1 Deregulated genes in *Suv420h2* ko splenic B cells (collaboration with Patrick Kremer laboratory; Gene Center, Munich)**

AffyID	SYMBOL	Fold change	Deregulation
10607950	NA	3.136436667	up
10340272	NA	1.614476667	up
10576807	Cd209d	1.437966667	up
10403006	NA	1.434246667	up
10549635	Lilra5	1.405286667	up
10474419	Lgr4	1.32121	up
10407387	Gm7120	1.302866667	up
10608711	Erdr1	1.293216667	up
10489721	IAP	1.28211	up
10528205	IAP	1.28211	up
10484355	IAP	1.28211	up
10363005	IAP	1.28211	up
10582983	IAP	1.28211	up
10512463	IAP	1.28211	up
10543029	IAP	1.28211	up
10528165	IAP	1.28211	up
10422247	IAP	1.28211	up
10345436	IAP	1.28211	up
10589974	IAP	1.28211	up
10527425	IAP	1.28211	up
10522742	IAP	1.28211	up
10500527	IAP	1.28211	up
10581009	IAP	1.28211	up
10568534	NA	1.28211	up
10496336	IAP	1.28211	up
10356762	IAP	1.28211	up
10476399	NA	1.28211	up
10607429	IAP	1.28211	up
10410927	IAP	1.28211	up
10374352	NA	1.28211	up
10457667	Gm4638	1.28211	up
10586863	IAP	1.28211	up
10416698	NA	1.28211	up
10468487	IAP	1.28211	up
10483161	NA	1.28211	up
10469127	IAP	1.28211	up
10603803	IAP	1.28211	up
10541129	IAP	1.28211	up
10574434	IAP	1.28211	up
10485355	Gm4638	1.28211	up
10346222	IAP	1.28211	up
10497327	IAP	1.28211	up
10374183	Gm4638	1.28211	up
10565598	IAP	1.28211	up
10395275	IAP	1.28211	up
10416696	NA	1.28211	up
10427389	IAP	1.28211	up
10603549	IAP	1.28211	up
10478746	IAP	1.28211	up
10496167	IAP	1.28211	up
10448230	IAP	1.28211	up
10540531	IAP	1.28211	up
10375324	IAP	1.28211	up
10363561	IAP	1.28211	up
10357298	IAP	1.28211	up
10359642	IAP	1.28211	up
10424377	Gm4638	1.28211	up
10574432	Gm4638	1.28211	up
10423971	Pkhd11	1.046233333	down
10338989	NA	1.059156667	down
10542172	Clec1b	1.059886667	down
10338966	NA	1.06161	down

## Appendix

10344207	NA	1.063043333	down
10342917	NA	1.063793333	down
10375055	F830116E18Rik	1.06413	down
10570894	Ank1	1.065766667	down
10338577	NA	1.070986667	down
10344331	NA	1.07137	down
10339167	NA	1.072073333	down
10507500	Slc6a9	1.07515	down
10560624	Apoe	1.07903	down
10343453	NA	1.080766667	down
10339089	NA	1.08169	down
10545086	Snca	1.084823333	down
10344100	NA	1.089166667	down
10341384	NA	1.091923333	down
10342786	NA	1.102373333	down
10342114	NA	1.104233333	down
10341669	NA	1.108736667	down
10362428	Trdn	1.110586667	down
10344113	NA	1.111903333	down
10459391	Fech	1.119923333	down
10473356	Ube2l6	1.126663333	down
10342748	NA	1.130133333	down
10538459	Aqp1	1.140773333	down
10481627	Lcn2	1.147273333	down
10504692	Tmod1	1.14755	down
10339273	NA	1.148993333	down
10473349	Ypel4	1.150043333	down
10342475	NA	1.166316667	down
10342827	NA	1.16659	down
10391649	Slc4a1	1.167036667	down
10400483	Slc25a21	1.201346667	down
10339166	NA	1.21424	down
10338617	NA	1.21689	down
10338332	NA	1.227113333	down
10466659	Gda	1.229463333	down
10567564	Cdr2	1.236206667	down
10515848	Ermap	1.239373333	down
10539080	St3gal5	1.240966667	down
10461979	Aldh1a1	1.24488	down
10339567	NA	1.252913333	down
10338462	NA	1.258173333	down
10602372	Alas2	1.262343333	down
10338546	NA	1.27335	down
10344068	NA	1.27363	down
10340865	NA	1.27596	down
10341388	NA	1.279896667	down
10537179	Bpgm	1.293256667	down
10429968	NA	1.29663	down
10344233	NA	1.30486	down
10474229	Cd59a	1.310603333	down
10341472	NA	1.33807	down
10342272	NA	1.339823333	down
10340845	NA	1.362546667	down
10338377	NA	1.36328	down
10549700	Suv420h2	1.369103333	down
10490923	Car2	1.370346667	down
10341301	NA	1.38235	down
10351905	Spna1	1.3849	down
10342791	NA	1.407726667	down
10339230	NA	1.415156667	down
10342166	NA	1.42899	down
10589535	Ngp	1.430423333	down
10607972	Kdm5d	1.430866667	down
10608001	Eif2s3y	1.431806667	down
10399710	Rsad2	1.43485	down
10340162	NA	1.44882	down
10344372	NA	1.450243333	down
10445046	Trim10	1.460053333	down
10544383	Kel	1.470013333	down
10534389	Cldn13	1.480493333	down
10608710	Ahsp	1.48214	down
10509002	Rhd	1.48334	down
10419151	Ear1	1.483993333	down
10608138	Ddx3y	1.52976	down

## Appendix

10338428	NA	1.534223333	down
10499062	Fhdc1	1.56721	down
10342665	NA	1.58273	down
10343713	NA	1.59116	down
10445192	Rhag	1.61186	down
10486664	Epb4.2	1.63213	down
10451670	Tspo2	1.633906667	down
10338259	NA	1.64597	down
10550332	Slc1a5	1.694356667	down
10573054	Gypa	1.7213	down
10338649	NA	1.776523333	down
10608107	Uty	1.779483333	down
10549655	Eps8l1	1.803306667	down
10561055	Ceacam2	3.841073333	down

## 5.2 Appendix (Results II)

Table 5.2 Deregulated genes in *Setdb1*<sup>Mb1</sup> pro B cells

Gene	Control	cko	Ratio	Regulated
Stfa1	0.0	3.2	3234.8	up
Amd2	0.0	9.9	3049.8	up
2410003L11Rik	0.0	2.5	2480.2	up
St6galnac5	0.0	3.9	191.7	up
Tubb3	2.0	211.7	106.4	up
Tdrd9	0.0	4.3	93.1	up
Mc1r	0.2	3.8	24.1	up
Acoxl	0.2	3.4	20.0	up
Cbr3	0.4	6.6	18.4	up
Nrg4	0.9	13.4	15.2	up
Gm5483	0.2	2.3	13.9	up
Def8	4.5	53.0	11.9	up
Lalba	8.2	85.7	10.4	up
Stfa3	0.5	5.1	10.3	up
BC100530	0.6	5.6	9.7	up
AI506816	63.3	599.7	9.5	up
Fcgr2b	12.4	95.4	7.7	up
Fbln1	0.3	2.6	7.4	up
Bglap	0.5	3.2	7.0	up
4921508A21Rik	3.4	23.1	6.8	up
Stfa2	0.5	3.2	6.7	up
Stag3	2.7	17.0	6.4	up
Rps4y2	19.9	120.3	6.1	up
Dmc1	0.6	3.3	6.0	up
Rims3	0.8	4.5	5.9	up
Lrrc32	0.5	3.0	5.9	up
2810474O19Rik	6.5	37.2	5.7	up
1700030C10Rik	1.3	7.4	5.6	up
Rpl39l	3.4	18.8	5.5	up
Akap5	1.0	5.3	5.5	up
Cd209c	1.3	7.1	5.4	up
Ly6k	5.2	27.9	5.3	up
Clec2g	0.6	2.9	5.2	up
Thy1	15.0	76.9	5.1	up
Cldn4	7.0	35.8	5.1	up
Tmem40	0.5	2.6	5.1	up
1700097N02Rik	20.8	90.9	4.4	up
Pnp2	5.7	24.2	4.3	up
Gstp2	1.4	5.8	4.2	up
Tmem90a	1.5	6.4	4.1	up
Apoc2	0.9	3.5	3.9	up
Csf1r	0.7	2.9	3.9	up
Fbxo17	2.4	9.3	3.8	up
H19	0.8	2.9	3.8	up
Myl9	1.1	3.9	3.7	up
Olfir1033	2.3	8.5	3.7	up
Rffl	7.3	26.8	3.7	up
2310039L15Rik	1.8	6.6	3.6	up
Gm12505	2.9	10.1	3.5	up
Trim24	10.4	36.4	3.5	up

## Appendix

Rbm44	0.9	3.1	3.5	up
BC048943	1.9	6.6	3.5	up
Gm5111	6.1	20.6	3.4	up
Serpina3f	5.3	17.5	3.3	up
A130040M12Rik	201.2	649.6	3.2	up
Apoe	19.3	62.2	3.2	up
Tnf	2.0	6.5	3.2	up
Lta	2.9	9.1	3.1	up
Vpreb2	242.1	729.5	3.0	up
Ephb2	2.6	7.9	3.0	up
Rnf187	228.3	665.5	2.9	up
Pglyrp2	1.8	5.3	2.9	up
Ckap4	7.0	20.3	2.9	up
S100a4	7.2	20.8	2.9	up
Cd244	2.3	6.4	2.8	up
Wdr92	8.7	23.9	2.8	up
Stfa2l1	1.3	3.5	2.7	up
Padi3	7.5	20.0	2.7	up
Ccl6	1.4	3.7	2.7	up
Prnp	4.5	12.0	2.7	up
3300005D01Rik	3.2	8.5	2.6	up
Snhg9	184.9	487.5	2.6	up
1700102H20Rik	2.3	5.9	2.6	up
Clcn5	2.0	5.4	2.6	up
Derl3	3.5	9.1	2.6	up
Ncs1	1.6	4.1	2.6	up
D10Wsu102e	12.3	31.4	2.6	up
Fcrla	65.2	166.2	2.5	up
BC024386	2.0	5.0	2.5	up
Cacna1h	1.4	3.6	2.5	up
Atp1b3	207.9	523.6	2.5	up
Cst7	3.6	9.1	2.5	up
Nr1d1	6.6	16.6	2.5	up
Flt3	6.6	16.6	2.5	up
Pde8a	1.9	4.7	2.5	up
Fgd2	6.9	17.3	2.5	up
Sema3g	2.9	7.3	2.5	up
Padi2	4.0	9.9	2.5	up
Socs2	16.4	40.1	2.5	up
Zcwpw1	6.4	15.6	2.4	up
Lst1	10.2	24.6	2.4	up
Sh2b2	2.5	6.0	2.4	up
Prr5	18.1	42.5	2.3	up
Fam69b	6.0	14.0	2.3	up
C030037D09Rik	2.9	6.7	2.3	up
0610010F05Rik	1.7	4.0	2.3	up
Gm16062	1.7	4.0	2.3	up
5730469M10Rik	3.5	7.9	2.3	up
Bcl2l11	6.6	15.3	2.3	up
Cyp4f18	55.6	126.5	2.3	up
Cebpb	3.1	7.1	2.3	up
Hcst	12.0	26.6	2.2	up
Pilra	2.0	4.3	2.2	up
Aig1	1.9	4.2	2.2	up
Ptk7	1.7	3.7	2.2	up
I730030J21Rik	3.0	6.5	2.2	up
Alyref2	22.4	49.0	2.2	up
Poli	3.4	7.4	2.2	up
Ston1	2.8	6.1	2.2	up
Vill	5.2	11.2	2.2	up
Fads3	2.3	5.0	2.1	up
Fanca	7.4	15.5	2.1	up
Tdrkh	2.6	5.6	2.1	up
Relt	6.0	12.5	2.1	up
1810044D09Rik	10.7	22.3	2.1	up
Dnahc8	3.1	6.4	2.1	up
Pnp	54.8	113.8	2.1	up
Ebpl	6.8	14.1	2.1	up
Iqcg	2.2	4.5	2.1	up
Cish	14.3	29.5	2.1	up
Aipl1	2.5	5.1	2.1	up
Deptor	2.1	4.3	2.1	up
Ahsa2	6.4	13.1	2.0	up
Ifitm6	4.5	9.2	2.0	up

## Appendix

Csf2rb	5.4	11.1	2.0	up
Grap2	4.4	8.9	2.0	up
Apobec2	3.8	7.8	2.0	up
2610524H06Rik	17.4	35.4	2.0	up
Bok	29.4	59.3	2.0	up
Myo10	2.7	5.4	2.0	up
Phlda3	4.6	2.3	0.5	down
Tcn2	17.3	8.6	0.5	down
Sik1	27.0	13.4	0.5	down
H2-T24	4.4	2.2	0.5	down
Rpl3l	10.6	5.2	0.5	down
Hspa12b	8.4	4.1	0.5	down
Tctn1	14.6	7.2	0.5	down
Calcoco1	20.6	10.1	0.5	down
4930562F07Rik	18.7	9.1	0.5	down
Irf7	9.2	4.5	0.5	down
Rtp4	8.3	4.0	0.5	down
Comt	99.4	48.2	0.5	down
Tubb2b	12.5	6.0	0.5	down
Btg1	52.5	25.4	0.5	down
Slc7a7	10.4	5.0	0.5	down
Hlx	8.2	3.9	0.5	down
2200002D01Rik	17.6	8.3	0.5	down
Hvcn1	41.9	19.7	0.5	down
Relb	9.5	4.4	0.5	down
Samhd1	14.3	6.6	0.5	down
2410066E13Rik	6.3	2.9	0.5	down
Bmf	23.6	10.9	0.5	down
Rapgef1	5.3	2.4	0.5	down
Nuak1	4.8	2.2	0.5	down
Arntl	63.5	29.0	0.5	down
Tcf7	10.1	4.6	0.5	down
Malat1	42.5	19.1	0.4	down
Tnk2	12.0	5.4	0.4	down
E230016M11Rik	5.1	2.3	0.4	down
Tbx6	10.2	4.5	0.4	down
Camk2b	15.6	6.9	0.4	down
Rab6b	4.7	2.1	0.4	down
Lmo4	65.9	29.1	0.4	down
1700027J07Rik	12.0	5.3	0.4	down
Zfp296	10.6	4.7	0.4	down
Ets2	4.9	2.1	0.4	down
Tcp11l2	15.6	6.8	0.4	down
Eya2	8.2	3.6	0.4	down
Otud1	8.5	3.7	0.4	down
Bcl6	4.6	2.0	0.4	down
Slit1	6.4	2.8	0.4	down
Crip2	14.7	6.3	0.4	down
Gchfr	4.1	1.8	0.4	down
Steap4	13.7	5.7	0.4	down
Apbb1	12.1	5.0	0.4	down
Chrb1	5.8	2.4	0.4	down
Cmtm8	6.1	2.5	0.4	down
Plekhg5	4.4	1.8	0.4	down
Emid1	14.9	6.1	0.4	down
Parvb	10.6	4.3	0.4	down
Stxbp1	5.7	2.3	0.4	down
Ankrd33b	14.9	6.0	0.4	down
Phlda1	3.4	1.4	0.4	down
Cass4	4.5	1.8	0.4	down
Capsl	7.3	2.9	0.4	down
Col27a1	11.1	4.3	0.4	down
Rab19	5.2	2.0	0.4	down
Slc2a8	4.3	1.6	0.4	down
Clec2i	3.3	1.2	0.4	down
Egr1	9.1	3.4	0.4	down
Mt1	34.2	12.8	0.4	down
Usp18	3.3	1.2	0.4	down
Mgst1	31.7	11.7	0.4	down
Cd55	3.7	1.4	0.4	down
Zbp1	7.7	2.8	0.4	down
Socs1	10.8	3.9	0.4	down
Neurl3	6.8	2.5	0.4	down
Ly6d	222.9	80.7	0.4	down



## Appendix

Rnase6	6.5	2.3	0.4	down
Ephx1	8.4	3.0	0.4	down
Kdm5b	6.7	2.3	0.4	down
Ccr7	9.2	3.2	0.3	down
Klrb1c	3.6	1.2	0.3	down
Hdac11	11.5	3.9	0.3	down
Gm16386	4.1	1.4	0.3	down
Neat1	21.4	7.2	0.3	down
Gpr155	3.4	1.1	0.3	down
Asb2	11.3	3.7	0.3	down
Rgs2	20.9	6.7	0.3	down
Amd1	10.2	3.3	0.3	down
Bcl3	4.6	1.5	0.3	down
Maf	10.4	3.3	0.3	down
Mt2	7.9	2.4	0.3	down
Enpep	5.3	1.6	0.3	down
H2-Q8	27.0	7.8	0.3	down
Klf2	47.7	13.8	0.3	down
Efh1d1	3.2	0.9	0.3	down
Dtx1	3.1	0.9	0.3	down
2010300C02Rik	15.6	4.4	0.3	down
H2-Eb1	11.7	3.2	0.3	down
Gh	3.5	1.0	0.3	down
Trib2	6.8	1.8	0.3	down
Cd72	133.3	34.6	0.3	down
Apobec1	16.6	4.2	0.3	down
Glcc1	7.7	1.9	0.2	down
Ntng2	3.3	0.8	0.2	down
Tmem163	3.7	0.9	0.2	down
Aim1	2.7	0.6	0.2	down
H2-Aa	16.0	3.5	0.2	down
Zap70	6.0	1.3	0.2	down
Kcna5	2.6	0.5	0.2	down
Rgs10	5.6	1.1	0.2	down
Nrgn	70.4	13.8	0.2	down
Srp3k	5.9	1.1	0.2	down
Inpp1	3.4	0.6	0.2	down
Hck	8.2	1.3	0.2	down
H2-Q9	3.3	0.5	0.2	down
Scd1	6.3	1.0	0.2	down
Cpm	9.7	1.5	0.2	down
Cd74	123.7	18.2	0.1	down
Cd38	5.3	0.8	0.1	down
Slc45a3	2.6	0.4	0.1	down
Heyl	16.8	2.2	0.1	down
Cd40	4.0	0.5	0.1	down
Sertad4	10.0	1.2	0.1	down
Speer2	7.1	0.9	0.1	down
Il2ra	2.5	0.3	0.1	down
H2-Q6	26.9	3.2	0.1	down
H2-Q7	20.9	2.3	0.1	down
Gas7	9.9	1.1	0.1	down
Serpina3n	2.3	0.2	0.1	down
Dlx1	3.0	0.3	0.1	down
Cacna1e	3.7	0.4	0.1	down
Fcrl6	10.0	0.9	0.1	down
Egr3	2.4	0.2	0.1	down
Ms4a1	7.5	0.7	0.1	down
Adam19	9.0	0.8	0.1	down
Reln	2.7	0.2	0.1	down
Faim3	17.8	1.1	0.1	down
Ikzf3	16.7	1.0	0.1	down
Polm	15.6	0.9	0.1	down
Col5a3	7.7	0.3	0.0	down
Fcer2a	2.8	0.1	0.0	down
Grb7	19.1	0.6	0.0	down
Cd2	10.0	0.3	0.0	down
Dlx1as	2.4	0.0	0.0	down

Table 5.3 Deregulated repetitive elements in *Setdb1*<sup>Mb1</sup> pro B cells

## Appendix

Repeat Family	Average Length	Copies	control reads	ko reads	control RPKM	ko RPKM	ratio
MuLV-int LTR ERV1	779.9	392	5333.621	82922.696	0.0184	0.2722	14.7585
RLTR3_MM LTR ERVK	382.3	31	227.593	1691.392	0.0202	0.1437	7.1133
MMTV-int LTR ERVK	980.5	35	275.819	1939.206	0.0090	0.0575	6.3634
RLTR6C_MM LTR ERV1	411.7	102	156.523	970.806	0.0047	0.0241	5.1018
MER49 LTR ERV1	84.6	7	0	2.406	0.0010	0.0051	5.0628
MMVL30-int LTR ERV1	720.5	956	4071.286	13848.719	0.0069	0.0211	3.0540
LTRIS4B LTR ERV1	317.5	71	0	45.713	0.0010	0.0030	3.0279
MER66-int LTR ERV1	117.7	23	5.076	16.842	0.0029	0.0072	2.5117
RLTR50A LTR ERVK	291.9	85	1.692	37.292	0.0011	0.0025	2.3432
RLTR4_MM LTR ERV1	466.6	239	3497.651	8070.801	0.0324	0.0734	2.2671
RLTR6_MM LTR ERV1	540	237	439.956	1049	0.0044	0.0092	2.0724
RNLTR23 LTR ERVK	99.4	5	1.692	3.609	0.0044	0.0083	1.8757
MER50 LTR ERV1	85.2	13	1.692	3.609	0.0025	0.0043	1.6847
Charlie17 DNA hAT-Charlie	233.2	134	16.921	40.901	0.0015	0.0023	1.4978
RLTR30D_RN LTR ERV1	278.3	64	285.125	434.276	0.0170	0.0254	1.4924
RLTR6-int LTR ERV1	1871.2	1220	1302.101	2929.259	0.0016	0.0023	1.4539
IAPLTR4 LTR ERVK	322.3	663	68.532	196.086	0.0013	0.0019	1.4520
MER95 LTR ERV1	204.5	54	0	4.812	0.0010	0.0014	1.4358
MADE2 DNA TcMar-Mariner	71.1	208	2.538	9.624	0.0012	0.0017	1.4090
ERV4_2-I_MM-int LTR ERVK	1530.8	313	28.766	233.378	0.0011	0.0015	1.4029
Eulor2A DNA? DNA?	149.2	41	0	2.406	0.0010	0.0014	1.3933
LTR86B2 LTR ERVL	184.8	36	1.692	4.812	0.0013	0.0017	1.3739
RNERVK23-int LTR ERVK	180.7	45	180.213	250.22	0.0232	0.0318	1.3717
MamRep1894 DNA hAT	92.2	36	0	1.203	0.0010	0.0014	1.3624
L1ME3D LINE L1	234.3	323	27.074	63.758	0.0014	0.0018	1.3570
IAPEY4_I-int LTR ERVK	808.6	521	178.521	388.563	0.0014	0.0019	1.3502
UCON12A Unknown Unknown	150	23	0	1.203	0.0010	0.0013	1.3487
RLTR4_MM-int LTR ERV1	1328.6	432	25123.183	34078.051	0.0448	0.0604	1.3485
X9_LINE LINE L1?	74.3	52	37.227	50.525	0.0106	0.0141	1.3236
Charlie16a DNA hAT-Charlie	163.2	197	20.306	32.48	0.0016	0.0020	1.2321
RMER16B3 LTR ERVK	260.8	101	0	6.015	0.0010	0.0012	1.2284
Charlie6 DNA hAT-Charlie	262.6	148	26.228	39.698	0.0017	0.0020	1.2069
RLTR45 LTR ERVK	451.2	1557	665.011	925.093	0.0019	0.0023	1.1902
MER136 DNA DNA	185.5	23	8.461	10.827	0.0030	0.0035	1.1859
MMERGLN-int LTR ERV1	1477.2	856	1192.112	1643.272	0.0019	0.0023	1.1837
MER106B DNA hAT-Charlie	183.4	242	13.537	24.06	0.0013	0.0015	1.1817
RLTR48B LTR ERV1	295.4	110	21.152	30.075	0.0017	0.0019	1.1663
Tigger8 DNA TcMar-Tigger	229	121	177.675	211.725	0.0074	0.0086	1.1658
RLTR31_MM LTR ERVK	377.4	145	21.152	33.683	0.0014	0.0016	1.1652
RLTR8 LTR ERVK	311	241	10.153	24.06	0.0011	0.0013	1.1634
MMETn-int LTR ERVK	1515.5	980	550.791	879.379	0.0014	0.0016	1.1614
Charlie18a DNA hAT-Charlie	153.7	242	41.457	54.134	0.0021	0.0025	1.1612
Chap1_MM DNA hAT-Charlie	149.8	153	32.151	40.901	0.0024	0.0028	1.1589
LTR86B1 LTR ERVL	174.8	49	5.076	7.218	0.0016	0.0018	1.1570
Kanga1b DNA TcMar-Tc2	203.1	121	0.846	4.812	0.0010	0.0012	1.1560
LTR16D2 LTR ERVL	166.2	65	1.692	3.609	0.0012	0.0013	1.1534
RLTR19-int LTR ERVK	335.5	2407	1195.496	1502.523	0.0025	0.0029	1.1533
LTRIS5 LTR ERV1	317.3	85	1.692	6.015	0.0011	0.0012	1.1508
RLTR20D LTR ERVK	265.4	777	65.147	105.862	0.0013	0.0015	1.1500
MER4B-int LTR ERV1	184.9	52	0.846	2.406	0.0011	0.0013	1.1491
UCON5 Unknown Unknown	144.7	56	0	1.203	0.0010	0.0011	1.1485
L1M2b LINE L1	127	16	18.614	21.654	0.0102	0.0117	1.1472
MER91C DNA hAT-Tip100	89.4	110	0.846	2.406	0.0011	0.0012	1.1461
ERV7_1-LTR_MM LTR ERVK	304.8	770	57.533	99.847	0.0012	0.0014	1.1448
Helitron3Na_MM RC Helitron	179.9	191	7.615	13.233	0.0012	0.0014	1.1338
L1M3f LINE L1	259.4	100	3.384	7.218	0.0011	0.0013	1.1307
IAPEY4_LTR LTR ERVK	313.8	95	5.076	9.624	0.0012	0.0013	1.1304
MER112 DNA hAT-Charlie	139.2	358	26.228	36.089	0.0015	0.0017	1.1296

## Appendix

MER97c DNA hAT-Tip100	297.8	64	0	2.406	0.0010	0.0011	1.1262
MMERGLN_LTR LTV ERV1	384	435	48.226	74.585	0.0013	0.0014	1.1224
RMER16C LTV ERVK	288.6	190	29.612	39.698	0.0015	0.0017	1.1194
MLTR18_MM LTV ERVK	348.2	427	18.614	38.495	0.0011	0.0013	1.1188
Tigger9b DNA TcMar-Tigger	217.9	360	6.769	16.842	0.0011	0.0012	1.1182
IAPLTR2a LTV ERVK	386.8	627	49.918	84.209	0.0012	0.0013	1.1173
LTR85c LTV Gypsy?	175.3	119	0	2.406	0.0010	0.0011	1.1153
ERVB4_1-I_MM-int LTV ERVK	1351.9	296	8.461	55.337	0.0010	0.0011	1.1147
LTR90B LTV LTV	151.6	91	0.846	2.406	0.0011	0.0012	1.1065
MER113B DNA hAT-Charlie	158.7	72	0	1.203	0.0010	0.0011	1.1053
Tigger9a DNA TcMar-Tigger	181.1	200	4.23	8.421	0.0011	0.0012	1.1036
MER54A LTV ERV1	272.9	273	5.076	13.233	0.0011	0.0012	1.1025
ERVB2_1-I_MM-int LTV ERVK	541.3	83	14.383	20.451	0.0013	0.0015	1.1023
RLTR31_Mur LTV ERVK	391.1	354	27.92	44.51	0.0012	0.0013	1.0997
MTB-int LTV ERV1-MaLR	698.8	592	204.749	265.859	0.0015	0.0016	1.0988
RLTR26C_MM LTV ERVK	364.8	454	61.763	84.209	0.0014	0.0015	1.0987
LIME5 LINE L1	271	146	0.846	4.812	0.0010	0.0011	1.0981
LTR85a LTV Gypsy?	160	234	0	3.609	0.0010	0.0011	1.0964
RLTRETN_Mm LTV ERVK	310	1355	118.45	169.62	0.0013	0.0014	1.0950
MamGyp-int LTV Gypsy	415.8	28	25.382	28.872	0.0032	0.0035	1.0943
MURVY-int LTV ERV1	3584.3	544	8.461	186.462	0.0010	0.0011	1.0909
Tigger16a DNA TcMar-Tigger	169.2	290	5.922	10.827	0.0011	0.0012	1.0892
MER90a LTV ERV1	229	581	10.153	22.857	0.0011	0.0012	1.0887
RLTR10F LTV ERVK	282.2	482	17.767	31.278	0.0011	0.0012	1.0879
L1MCa LINE L1	276.3	3565	346.042	458.336	0.0014	0.0015	1.0844
RLTR13F LTV ERVK	602.6	98	10.999	16.842	0.0012	0.0013	1.0834
MER73 LTV ERV1	239.5	135	0.846	3.609	0.0010	0.0011	1.0833
RLTR1C LTV ERV1	447.1	697	19.46	46.916	0.0011	0.0012	1.0829
RLTR9A2 LTV ERVK	313.6	227	6.769	13.233	0.0011	0.0012	1.0829
MER126 DNA DNA	155.2	118	0.846	2.406	0.0010	0.0011	1.0814
RLTR46B LTV ERVK	337	73	9.307	12.03	0.0014	0.0015	1.0803
RLTR35B_MM LTV ERV1	262.2	492	16.075	27.669	0.0011	0.0012	1.0799
LTR87 LTV ERV1?	191.6	214	2.538	6.015	0.0011	0.0011	1.0799
Tigger17 DNA TcMar-Tigger	169.5	315	22.844	28.872	0.0014	0.0015	1.0791
RLTR20A1 LTV ERVK	238.3	98	5.076	7.218	0.0012	0.0013	1.0753
MER45R DNA hAT-Tip100	180.3	375	4.23	9.624	0.0011	0.0011	1.0751
RLTR30D_MM LTV ERV1	344.4	268	5.922	13.233	0.0011	0.0011	1.0744
RLTR1B LTV ERV1	418.2	893	34.689	64.961	0.0011	0.0012	1.0742
RLTR12H LTV ERVK	341.4	188	6.769	12.03	0.0011	0.0012	1.0741
Tigger18a DNA TcMar-Tigger	241.1	240	1.692	6.015	0.0010	0.0011	1.0726
RMER17A-int LTV ERVK	365.7	409	13.537	25.263	0.0011	0.0012	1.0719
MER44A DNA TcMar-Tigger	211.7	661	52.456	66.164	0.0014	0.0015	1.0713
ERVB4_2-LTR_MM LTV ERVK	468.8	144	13.537	19.248	0.0012	0.0013	1.0705
MER88 LTV ERV1	219	58	2.538	3.609	0.0012	0.0013	1.0703
RLTR1A2_MM LTV ERV1	422.1	406	20.306	33.683	0.0011	0.0012	1.0698
MER102a DNA hAT-Charlie	173.9	386	71.07	80.6	0.0021	0.0022	1.0690
Charlie21a DNA hAT-Charlie	250.5	174	1.692	4.812	0.0010	0.0011	1.0689
MER77B LTV ERV1	247.5	505	10.153	19.248	0.0011	0.0012	1.0673
LTR77 LTV ERV1	117.8	38	0.846	1.203	0.0012	0.0013	1.0671
RLTR12BD_Mm LTV ERVK	293.5	883	96.452	120.298	0.0014	0.0015	1.0671
MMERVK10C-int LTV ERVK	879.3	3230	153.139	350.068	0.0011	0.0011	1.0658
L4_A_Mam LINE RTE-X	186.8	630	17.767	26.466	0.0012	0.0012	1.0642
LTR78B LTV ERV1	175.7	405	22.844	28.872	0.0013	0.0014	1.0641
RLTR43C LTV ERVK	273.9	264	16.075	21.654	0.0012	0.0013	1.0631
ERVB4_3-LTR_MM LTV ERVK	424.9	137	0	3.609	0.0010	0.0011	1.0620
RLTR19A LTV ERVK	292	153	10.999	14.436	0.0012	0.0013	1.0617
AmnSINE1 SINE Deu	96.2	1311	111.681	126.313	0.0019	0.0020	1.0615
RLTR30E_MM LTV ERV1	364.4	320	16.075	24.06	0.0011	0.0012	1.0602
L1M8 LINE L1	251.3	201	1.692	4.812	0.0010	0.0011	1.0598
MLT1M LTV ERV1-MaLR	183.8	212	29.612	33.683	0.0018	0.0019	1.0594

## Appendix

RMER16B2 LTR ERVK	282.4	329	8.461	14.436	0.0011	0.0012	1.0589
MER102b DNA hAT-Charlie	179.2	507	27.92	34.886	0.0013	0.0014	1.0586
ETnERV-int LTR ERVK	1280.3	906	226.747	307.963	0.0012	0.0013	1.0586
RLTR23 LTR ERV1	323	1524	69.378	102.253	0.0011	0.0012	1.0585
MER58C DNA hAT-Charlie	111.4	436	14.383	18.045	0.0013	0.0014	1.0582
RLTR13A2 LTR ERVK	865.6	151	0.846	8.421	0.0010	0.0011	1.0576
RMER17A LTR ERVK	479.3	1819	54.995	108.268	0.0011	0.0011	1.0575
MLT1G1 LTR ERV-L-MaLR	215.4	1080	47.38	62.555	0.0012	0.0013	1.0542
RLTR30 LTR ERV1	258.6	390	1.692	7.218	0.0010	0.0011	1.0539
IAPLTR2_Mm LTR ERVK	450.4	1388	478.029	536.53	0.0018	0.0019	1.0530
MamTip2 DNA hAT-Tip100	98.8	892	32.151	38.495	0.0014	0.0014	1.0527
MLTR18C_MM LTR ERVK	437.5	468	19.46	31.278	0.0011	0.0012	1.0527
RLTR12G LTR ERVK	386.4	238	16.075	21.654	0.0012	0.0012	1.0516
RLTR9E LTR ERVK	353.5	1318	16.075	40.901	0.0010	0.0011	1.0515
L1M3d LINE L1	248.1	171	2.538	4.812	0.0011	0.0011	1.0506
MER96B DNA hAT-Tip100	151.3	410	5.076	8.421	0.0011	0.0011	1.0498
MER91B DNA hAT-Tip100	96.6	229	25.382	27.669	0.0021	0.0023	1.0481
ERV3-16A3_LTR LTR ERV-L	205.8	279	0.846	3.609	0.0010	0.0011	1.0474
MLT1K LTR ERV-L-MaLR	181.5	2759	744.541	803.591	0.0025	0.0026	1.0474
MLT1-int LTR ERV-L-MaLR	234.6	285	8.461	12.03	0.0011	0.0012	1.0474
LTRIS4A LTR ERV1	347	228	6.769	10.827	0.0011	0.0011	1.0472
MamRep1527 LTR LTR	157.9	372	0.846	3.609	0.0010	0.0011	1.0464
ERVB4_1C-LTR_Mm LTR ERVK	224.9	961	45.688	57.743	0.0012	0.0013	1.0460
RLTR1F_Mm LTR ERVK	412.1	83	0.846	2.406	0.0010	0.0011	1.0445
L1MC4a LINE L1	232.4	3077	245.36	287.513	0.0013	0.0014	1.0439
MER34A LTR ERV1	199	694	18.614	25.263	0.0011	0.0012	1.0424
MER119 DNA hAT-Charlie	212.2	404	12.691	16.842	0.0011	0.0012	1.0422
RLTR1B-int LTR ERV1	707.3	798	65.147	91.427	0.0011	0.0012	1.0417
BGLII_A LTR ERVK	316.7	292	10.153	14.436	0.0011	0.0012	1.0417
L1M3de LINE L1	282.8	157	1.692	3.609	0.0010	0.0011	1.0416
RLTR20B1 LTR ERVK	310.9	668	32.151	42.104	0.0012	0.0012	1.0415
UCON29 DNA PiggyBac?	166.9	141	2.538	3.609	0.0011	0.0012	1.0411
RLTR28B LTR ERV-L	279.5	428	25.382	31.278	0.0012	0.0013	1.0407
LTR16B1 LTR ERV-L	231.7	299	4.23	7.218	0.0011	0.0011	1.0406
MER52-int LTR ERV1	233.9	206	5.076	7.218	0.0011	0.0011	1.0402
BGLII_Mur LTR ERVK	340.8	611	18.614	27.669	0.0011	0.0011	1.0399
LTR37B LTR ERV1	199.9	386	1.692	4.812	0.0010	0.0011	1.0396
RLTR19C LTR ERVK	333.4	248	17.767	21.654	0.0012	0.0013	1.0387
RLTR13D2 LTR ERVK	730.5	270	10.153	18.045	0.0011	0.0011	1.0381
MLT1A1 LTR ERV-L-MaLR	154.3	9015	143.832	200.898	0.0011	0.0011	1.0372
Tigger16b DNA TcMar-Tigger	132.5	197	43.15	45.713	0.0027	0.0028	1.0370
ORR1D1-int LTR ERV-L-MaLR	466	1715	108.297	140.749	0.0011	0.0012	1.0358
MamRep564 Unknown Unknown	147	125	1.692	2.406	0.0011	0.0011	1.0356
MT-int LTR ERV-L-MaLR	401.5	438	25.382	32.48	0.0011	0.0012	1.0353
RLTR10D LTR ERVK	342	491	11.845	18.045	0.0011	0.0011	1.0345
MER74A LTR ERV-L	244.2	731	21.998	28.872	0.0011	0.0012	1.0343
RLTR51B_Mm LTR ERVK?	242	662	10.999	16.842	0.0011	0.0011	1.0341
IAPLTR1_Mm LTR ERVK	368.5	1492	268.204	295.933	0.0015	0.0015	1.0339
Charlie24 DNA hAT-Charlie	213.1	374	43.996	48.119	0.0016	0.0016	1.0333
MER67B LTR ERV1	224.5	220	4.23	6.015	0.0011	0.0011	1.0333
MLT2B4 LTR ERV-L	220.8	1871	13.537	27.669	0.0010	0.0011	1.0331
LTR79 LTR ERV-L	165.8	380	8.461	10.827	0.0011	0.0012	1.0331
RLTR27 LTR ERVK	168.8	1621	13.537	22.857	0.0010	0.0011	1.0325
L1M3b LINE L1	286.2	169	0.846	2.406	0.0010	0.0010	1.0317
MER58B DNA hAT-Charlie	164.3	2843	87.991	104.659	0.0012	0.0012	1.0300
MER99 DNA hAT?	224.7	198	3.384	4.812	0.0011	0.0011	1.0298
RLTR22_Mus LTR ERVK	448.8	1191	73.608	91.427	0.0011	0.0012	1.0293
MER68B LTR ERV-L	220.5	237	16.075	18.045	0.0013	0.0013	1.0288
LTR16B LTR ERV-L	213.7	260	10.153	12.03	0.0012	0.0012	1.0286
LTR55 LTR ERV-L?	265.5	160	0	1.203	0.0010	0.0010	1.0283

## Appendix

L1M3c LINE L1	288.7	337	0.846	3.609	0.0010	0.0010	1.0282
RMER6-int LTR ERVK	425.5	607	11.845	19.248	0.0010	0.0011	1.0274
LTR72_RN LTR ERV1	219.6	1000	60.917	68.57	0.0013	0.0013	1.0273
IAPLTR3-int LTR ERVK	966.3	1773	63.455	111.877	0.0010	0.0011	1.0273
Arthur1 DNA hAT-Tip100	273.1	920	28.766	36.089	0.0011	0.0011	1.0262
RMER16A2 LTR ERVK	264.4	273	16.921	19.248	0.0012	0.0013	1.0261
Charlie2a DNA hAT-Charlie	266.9	505	11.845	15.639	0.0011	0.0011	1.0259
LTR75B LTR ERV1	195.6	69	0.846	1.203	0.0011	0.0011	1.0249
L1MEa LINE L1	241.7	169	2.538	3.609	0.0011	0.0011	1.0247
MLTR11A LTR ERVK	327.5	2124	60.071	78.194	0.0011	0.0011	1.0240
RLTR15 LTR ERVK	274.5	2437	124.372	143.155	0.0012	0.0012	1.0237
L1MDb LINE L1	241.3	583	35.535	39.698	0.0013	0.0013	1.0236
Plat_L3 LINE CR1	153.2	579	20.306	22.857	0.0012	0.0013	1.0234
MLTR18A_MM LTR ERVK	365.8	299	9.307	12.03	0.0011	0.0011	1.0229
MER5B DNA hAT-Charlie	107.3	5140	106.605	121.501	0.0012	0.0012	1.0226
ERV4_3-I_MM-int LTR ERVK	1534.1	99	2.538	6.015	0.0010	0.0010	1.0225
RLTR41 LTR ERV1	339	1121	15.229	24.06	0.0010	0.0011	1.0223
MLT2B2 LTR ERV1	226.3	1351	16.921	24.06	0.0011	0.0011	1.0221
MLT1A0-int LTR ERV1-MaLR	340.2	206	0.846	2.406	0.0010	0.0010	1.0220
L1ME4a LINE L1	207.1	1560	104.913	114.283	0.0013	0.0014	1.0219
MamGypLTR2b LTR Gypsy	142.8	109	0.846	1.203	0.0011	0.0011	1.0218
MER70B LTR ERV1	229.4	254	0	1.203	0.0010	0.0010	1.0206
ORR1A1-int LTR ERV1-MaLR	919.5	859	73.608	91.427	0.0011	0.0011	1.0206
Charlie2b DNA hAT-Charlie	236.1	884	34.689	39.698	0.0012	0.0012	1.0206
MLT1H1-int LTR ERV1-MaLR	341.4	49	0.846	1.203	0.0011	0.0011	1.0203
RLTR10B2 LTR ERVK	332.5	854	26.228	32.48	0.0011	0.0011	1.0202
RLTR9D2 LTR ERVK	321.8	194	27.074	28.872	0.0014	0.0015	1.0201
MER44C DNA TcMar-Tigger	275.3	185	8.461	9.624	0.0012	0.0012	1.0196
RLTR31A_MM LTR ERVK	369	209	12.691	14.436	0.0012	0.0012	1.0194
MERVK26-int LTR ERVK	628.1	1001	17.767	30.075	0.0010	0.0010	1.0190
RMER6D LTR ERVK	332.6	2411	79.531	96.238	0.0011	0.0011	1.0190
X7B_LINE LINE CR1	104.7	162	6.769	7.218	0.0014	0.0014	1.0189
RMER21A LTR ERV1	355.5	622	150.6	157.591	0.0017	0.0017	1.0188
MuRRS4-int LTR ERV1	881.3	2735	109.989	156.388	0.0010	0.0011	1.0184
RMER20C_MM LTR ERVK	327.1	1595	53.302	63.758	0.0011	0.0011	1.0182
L1ME3C LINE L1	235.3	326	3.384	4.812	0.0010	0.0011	1.0178
MER76 LTR ERV1	228.4	247	20.306	21.654	0.0014	0.0014	1.0176
RLTR26D_MM LTR ERVK	395.2	247	4.23	6.015	0.0010	0.0011	1.0175
RLTR11B LTR ERVK	317.1	1049	260.589	270.671	0.0018	0.0018	1.0170
RMER17D LTR ERVK	400	1166	49.072	57.743	0.0011	0.0011	1.0168
LTR40b LTR ERV1	217.7	577	5.076	7.218	0.0010	0.0011	1.0164
Tigger19a DNA TcMar-Tigger	125.3	771	24.536	26.466	0.0013	0.0013	1.0159
RLTR11A LTR ERVK	353.7	2655	350.273	370.518	0.0014	0.0014	1.0157
RLTR19 LTR ERVK	303.2	944	28.766	33.683	0.0011	0.0011	1.0156
Charlie29a DNA hAT-Charlie	211.1	324	8.461	9.624	0.0011	0.0011	1.0151
Kanga1c DNA TcMar-Tc2	193.9	184	4.23	4.812	0.0011	0.0011	1.0146
MER74C LTR ERV1	215	114	0.846	1.203	0.0010	0.0010	1.0141
RLTR13B2 LTR ERVK	716	314	7.615	10.827	0.0010	0.0010	1.0138
ERV5_1-I_MM-int LTR ERVK	699.1	292	6.769	9.624	0.0010	0.0010	1.0135
MERV1_LTR LTR ERV1	321.2	615	32.997	36.089	0.0012	0.0012	1.0134
MLT1H-int LTR ERV1-MaLR	243.7	379	11.845	13.233	0.0011	0.0011	1.0133
RLTR6B_MM LTR ERV1	541.9	646	82.069	87.818	0.0012	0.0013	1.0133
RLTR34B_MM LTR ERVK	366.6	527	35.535	38.495	0.0012	0.0012	1.0129
RLTR44-int LTR ERVK	804.6	653	63.455	70.976	0.0011	0.0011	1.0128
ORR1A4-int LTR ERV1-MaLR	773	1326	120.142	134.734	0.0011	0.0011	1.0127
IAPEY3C_LTR LTR ERVK	288.9	475	10.153	12.03	0.0011	0.0011	1.0127
MamGypLTR3a LTR Gypsy	170.6	161	0.846	1.203	0.0010	0.0010	1.0126
RLTR9D LTR ERVK	328.5	803	16.921	20.451	0.0011	0.0011	1.0126
ORR1D2 LTR ERV1-MaLR	241.9	15061	440.802	492.019	0.0011	0.0011	1.0125
MLT2B5 LTR ERV1	191	442	8.461	9.624	0.0011	0.0011	1.0125

## Appendix

MER20B DNA hAT-Charlie	202.2	831	93.068	96.238	0.0016	0.0016	1.0121
L4_C_Mam LINE RTE-X	173.9	590	5.922	7.218	0.0011	0.0011	1.0119
L1M4a2 LINE L1	242.4	736	10.999	13.233	0.0011	0.0011	1.0118
ERV1-int LTR ERV1	316.7	721	3.384	6.015	0.0010	0.0010	1.0114
MLT1H1 LTR ERV1-MaLR	209.7	1001	8.461	10.827	0.0010	0.0011	1.0108
MER67D LTR ERV1	238.6	468	0	1.203	0.0010	0.0010	1.0108
IAPEY_LTR LTR ERV1	337.1	1410	27.074	32.48	0.0011	0.0011	1.0108
RLTR10C LTR ERV1	414.6	1779	142.14	151.576	0.0012	0.0012	1.0107
L1_Mus2 LINE L1	756.3	21079	2400.299	2592.424	0.0012	0.0012	1.0105
HAL1 LINE L1	225.9	3550	121.834	131.125	0.0012	0.0012	1.0101
RLTR33 LTR ERV1	469.5	1336	58.379	64.961	0.0011	0.0011	1.0096
L1MB1 LINE L1	268	2524	32.997	39.698	0.0010	0.0011	1.0094
RLTR13A3 LTR ERV1	847.3	243	7.615	9.624	0.0010	0.0010	1.0094
RLTR44C LTR ERV1	395.6	198	54.148	55.337	0.0017	0.0017	1.0090
L1VL1 LINE L1	1106.4	1797	33.843	51.728	0.0010	0.0010	1.0088
RLTR19A2 LTR ERV1	235.4	94	3.384	3.609	0.0012	0.0012	1.0088
ERV1-E-int LTR ERV1	255.1	1044	28.766	31.278	0.0011	0.0011	1.0085
LTR102_Mam LTR ERV1	217.3	215	0.846	1.203	0.0010	0.0010	1.0075
RMER16-int LTR ERV1	653.8	2599	56.687	69.773	0.0010	0.0010	1.0075
MLT1A1-int LTR ERV1-MaLR	362.1	130	0.846	1.203	0.0010	0.0010	1.0075
L1_Mus1 LINE L1	807.3	27252	2248.852	2415.586	0.0011	0.0011	1.0069
CR1_Mam LINE CR1	177.1	257	15.229	15.639	0.0013	0.0013	1.0067
RLTR13E LTR ERV1	616.5	603	26.228	28.872	0.0011	0.0011	1.0066
RLTR20B5_MM LTR ERV1	327.8	868	36.381	38.495	0.0011	0.0011	1.0066
IAP-d-int LTR ERV1	486.1	1323	12.691	16.842	0.0010	0.0010	1.0063
IAPLTR1a_Mm LTR ERV1	337.4	1786	72.762	76.991	0.0011	0.0011	1.0063
ERVB4_1B-I_MM-int LTR ERV1	549.8	1368	10.999	15.639	0.0010	0.0010	1.0061
MERV1_I-int LTR ERV1	701	1365	91.375	97.441	0.0011	0.0011	1.0058
MLT2D LTR ERV1	213.9	1567	16.075	18.045	0.0010	0.0011	1.0056
L1_Rod LINE L1	372.8	11315	134.525	158.794	0.0010	0.0010	1.0056
HAL1ME LINE L1	197.9	900	10.999	12.03	0.0011	0.0011	1.0055
Tigger5 DNA TcMar-Tigger	183.4	1975	85.453	87.818	0.0012	0.0012	1.0053
RLTR53_MM LTR ERV1	285.5	481	120.142	121.501	0.0019	0.0019	1.0053
ETnERV3-int LTR ERV1	1382.9	560	12.691	16.842	0.0010	0.0010	1.0053
MER121 DNA TcMar?	195.7	588	4.23	4.812	0.0010	0.0010	1.0049
RLTR14-int LTR ERV1	440.6	1816	101.528	105.862	0.0011	0.0011	1.0048
MER58A DNA hAT-Charlie	155.4	3714	93.068	96.238	0.0012	0.0012	1.0047
MER34 LTR ERV1	164	737	4.23	4.812	0.0010	0.0010	1.0047
BGLII_Mus LTR ERV1	298.7	615	13.537	14.436	0.0011	0.0011	1.0046
L1Md_F LINE L1	1472.9	4016	119.296	146.764	0.0010	0.0010	1.0046
MER115 DNA hAT-Tip100	193.9	521	84.607	85.412	0.0018	0.0018	1.0043
MLT1D-int LTR ERV1-MaLR	302.1	176	3.384	3.609	0.0011	0.0011	1.0040
MLTR31D_MM LTR ERV1	418.4	318	12.691	13.233	0.0011	0.0011	1.0037
Charlie5 DNA hAT-Charlie	178.3	1547	313.046	315.181	0.0021	0.0021	1.0036
IAPLTR4_I LTR ERV1	557.9	183	0.846	1.203	0.0010	0.0010	1.0035
L1M6 LINE L1	189.9	329	3.384	3.609	0.0011	0.0011	1.0034
MLT1A LTR ERV1-MaLR	190	4950	66.839	69.773	0.0011	0.0011	1.0029
MYSERV6-int LTR ERV1	549	3183	112.527	117.892	0.0011	0.0011	1.0029
Lx2A LINE L1	601.5	1364	10.999	13.233	0.0010	0.0010	1.0027
MER92-int LTR ERV1	256.1	324	3.384	3.609	0.0010	0.0010	1.0026
L1MDa LINE L1	271.4	3451	68.532	70.976	0.0011	0.0011	1.0024
MMERV10D3_I-int LTR ERV1	1077.1	458	10.999	12.03	0.0010	0.0010	1.0020
MamRep1879 DNA hAT-Tip100?	125.8	105	14.383	14.436	0.0021	0.0021	1.0019
ORSL-2b DNA hAT-Tip100	222.8	210	5.922	6.015	0.0011	0.0011	1.0018
MuRRS-int LTR ERV1	1307.4	996	39.765	42.104	0.0010	0.0010	1.0017
L1MCb LINE L1	287.1	713	0.846	1.203	0.0010	0.0010	1.0017
Lx2A1 LINE L1	480.8	2739	25.382	27.669	0.0010	0.0010	1.0017
Looper DNA PiggyBac	238.6	209	5.922	6.015	0.0011	0.0011	1.0017
RLTR44A LTR ERV1	296.3	259	20.306	20.451	0.0013	0.0013	1.0015
Kanga2_a DNA TcMar-Tc2	226.8	356	20.306	20.451	0.0013	0.0013	1.0014



## Appendix

RMER1A Other Other	542.9	3854	458.57	461.945	0.0012	0.0012	1.0013
MMERVK9E_I-int LTR ERVVK	918.9	1568	84.607	86.615	0.0011	0.0011	1.0013
RMER17C-int LTR ERVVK	381.9	1940	79.531	80.6	0.0011	0.0011	1.0013
RLTR5_MM LTR ERV1	605.8	601	6.769	7.218	0.0010	0.0010	1.0012
L1MA4A LINE L1	293.2	2709	50.764	51.728	0.0011	0.0011	1.0011
AmnSINE2 SINE Deu	89.3	391	14.383	14.436	0.0014	0.0014	1.0011
MLT1F1 LTR ERVVL-MaLR	212	1608	32.151	32.48	0.0011	0.0011	1.0009
ERVB7_3-LTR_MM LTR ERVVK	334.8	319	5.922	6.015	0.0011	0.0011	1.0008
RMER16A3 LTR ERVVK	154.9	699	5.922	6.015	0.0011	0.0011	1.0008
MLT1J-int LTR ERVVL-MaLR	233.7	506	5.922	6.015	0.0011	0.0011	1.0007
MLT2B1 LTR ERVVL	245.2	1637	17.767	18.045	0.0010	0.0010	1.0007
RLTR50B LTR ERVVK	438.6	334	5.922	6.015	0.0010	0.0010	1.0006
RLTR31M LTR ERVVK	376.9	264	60.071	60.149	0.0016	0.0016	1.0005
MusHAL1 LINE L1	587.9	3890	56.687	57.743	0.0010	0.0010	1.0005
IAPEY3-int LTR ERVVK	892	2345	32.997	33.683	0.0010	0.0010	1.0003
L1VL2 LINE L1	1127.6	4697	152.292	153.982	0.0010	0.0010	1.0003
MER92D LTR ERV1	178.7	70	0	0	0.0010	0.0010	1.0000
LTR90A LTR LTR	150.3	99	0	0	0.0010	0.0010	1.0000
UCON1 Unknown Unknown	128.5	15	0	0	0.0010	0.0010	1.0000
C573_MM Other Other	299.2	1738	0	0	0.0010	0.0010	1.0000
RLTR47_MM LTR ERV1	308	45	0	0	0.0010	0.0010	1.0000
MARINER1_EC DNA TcMar-Tigger	76.5	2	0	0	0.0010	0.0010	1.0000
MLT1E2-int LTR ERVVL-MaLR	292.4	35	0	0	0.0010	0.0010	1.0000
HY5 scRNA scRNA	74	1	0	0	0.0010	0.0010	1.0000
ORSL-2a DNA hAT-Tip100	194.4	69	0	0	0.0010	0.0010	1.0000
nhAT5a_ML DNA hAT-Charlie	100	4	0	0	0.0010	0.0010	1.0000
Tigger17b DNA TcMar-Tigger	213.5	72	0	0	0.0010	0.0010	1.0000
L1P5 LINE L1	135.6	166	0	0	0.0010	0.0010	1.0000
Ricksha_a DNA MULE-MuDR	265.7	30	0	0	0.0010	0.0010	1.0000
Eulor10 DNA? DNA?	134.4	29	0	0	0.0010	0.0010	1.0000
MamGypLTR1d LTR Gypsy	168	116	0	0	0.0010	0.0010	1.0000
MER133B DNA? DNA?	92.3	22	0	0	0.0010	0.0010	1.0000
Helitron1Nb_Mam RC Helitron	194.2	92	0	0	0.0010	0.0010	1.0000
LTR16D LTR ERVVL	196.8	185	0	0	0.0010	0.0010	1.0000
UCON28a Unknown Unknown	177.9	87	0	0	0.0010	0.0010	1.0000
MER72 LTR ERV1	100.4	9	0	0	0.0010	0.0010	1.0000
MER66A LTR ERV1	87.6	5	0	0	0.0010	0.0010	1.0000
UCON12 Unknown Unknown	136	25	0	0	0.0010	0.0010	1.0000
Cheshire_Mars_ DNA hAT-Charlie	249.3	6	0	0	0.0010	0.0010	1.0000
MER47C DNA TcMar-Tigger	76.5	14	0	0	0.0010	0.0010	1.0000
Eulor3 DNA? DNA?	161.6	28	0	0	0.0010	0.0010	1.0000
UCON25 Unknown Unknown	91.4	23	0	0	0.0010	0.0010	1.0000
L1M6B LINE L1	140	43	0	0	0.0010	0.0010	1.0000
UCON28b Unknown Unknown	213.2	49	0	0	0.0010	0.0010	1.0000
MER57-int LTR ERV1	135.7	202	0	0	0.0010	0.0010	1.0000
LTR27C LTR ERV1	119.9	7	0	0	0.0010	0.0010	1.0000
LTR45C LTR ERV1	115.2	24	0	0	0.0010	0.0010	1.0000
LTR39 LTR ERV1	87.5	11	0	0	0.0010	0.0010	1.0000
X1_LINE LINE CR1	234.9	40	0	0	0.0010	0.0010	1.0000
RLTR1 LTR ERV1	227.6	62	0	0	0.0010	0.0010	1.0000
LTR108c_Mam LTR ERVVL	232.5	15	0	0	0.0010	0.0010	1.0000
MER94B DNA hAT-Blackjack	94	86	0	0	0.0010	0.0010	1.0000
LTR9D LTR ERV1	77.2	11	0	0	0.0010	0.0010	1.0000
Eulor11 DNA DNA	171.8	49	0	0	0.0010	0.0010	1.0000
Charlie1b_Mars DNA hAT-Charlie	359.5	2	0	0	0.0010	0.0010	1.0000
MER110-int LTR ERV1	233.5	38	0	0	0.0010	0.0010	1.0000
UCON10 Unknown Unknown	178	49	0	0	0.0010	0.0010	1.0000
X5A_LINE LINE CR1	106.6	69	0	0	0.0010	0.0010	1.0000
RMER3D2 LTR ERVVK	82.3	3	0	0	0.0010	0.0010	1.0000

## Appendix

MER133A DNA? DNA?	78.5	35	0	0	0.0010	0.0010	1.0000
LTR75_1 LTR ERV1	89.2	9	0	0	0.0010	0.0010	1.0000
MLT1K-int LTR ERV1-MaLR	269.2	8	0	0	0.0010	0.0010	1.0000
L5 LINE RTX-X	180.4	122	0	0	0.0010	0.0010	1.0000
UCON17 Unknown Unknown	198	23	0	0	0.0010	0.0010	1.0000
Eulor9C DNA DNA	96.9	71	0	0	0.0010	0.0010	1.0000
UCON7 DNA? DNA?	166.4	99	0	0	0.0010	0.0010	1.0000
Charlie22a DNA hAT-Charlie	175.3	163	0	0	0.0010	0.0010	1.0000
HERVL32-int LTR ERV1	124	1	0	0	0.0010	0.0010	1.0000
LTR9A1 LTR ERV1	114.6	11	0	0	0.0010	0.0010	1.0000
L1MEg1 LINE L1	206.1	92	0	0	0.0010	0.0010	1.0000
X2_LINE LINE CR1	141.4	32	0	0	0.0010	0.0010	1.0000
MER57E3 LTR ERV1	187.5	4	0	0	0.0010	0.0010	1.0000
MLT1L-int LTR ERV1-MaLR	211	13	0	0	0.0010	0.0010	1.0000
MER132 DNA TcMar-Pogo	127.6	21	0	0	0.0010	0.0010	1.0000
RLTR46 LTR ERV1	256.2	46	0	0	0.0010	0.0010	1.0000
LTR68 LTR ERV1	168.5	106	0	0	0.0010	0.0010	1.0000
Eulor6A DNA? DNA?	122.6	21	0	0	0.0010	0.0010	1.0000
Helitron2Na_Mam RC? Helitron?	147.4	74	0	0	0.0010	0.0010	1.0000
UCON19 Unknown Unknown	153.8	24	0	0	0.0010	0.0010	1.0000
MER83A-int LTR ERV1	119	1	0	0	0.0010	0.0010	1.0000
MLT1H2-int LTR ERV1-MaLR	342.6	53	0	0	0.0010	0.0010	1.0000
Ricksha DNA MULE-MuDR	288.3	60	0	0	0.0010	0.0010	1.0000
LTR52-int LTR ERV1	317	50	0	0	0.0010	0.0010	1.0000
MamRep488 DNA hAT-Tip100	121	29	0	0	0.0010	0.0010	1.0000
MLT1J1-int LTR ERV1-MaLR	258.1	67	0	0	0.0010	0.0010	1.0000
LTR106_Mam LTR LTR	229.5	164	0	0	0.0010	0.0010	1.0000
RMER3D3 LTR ERV1	95.6	8	0	0	0.0010	0.0010	1.0000
Eulor6B DNA? DNA?	162.8	56	0	0	0.0010	0.0010	1.0000
MLT1F1-int LTR ERV1-MaLR	265.9	89	0	0	0.0010	0.0010	1.0000
Tigger11a DNA TcMar-Tigger	211.1	75	0	0	0.0010	0.0010	1.0000
Charlie7a DNA hAT-Charlie	145.5	102	0	0	0.0010	0.0010	1.0000
LTR34 LTR ERV1	115.5	8	0	0	0.0010	0.0010	1.0000
UCON15 Unknown Unknown	160.8	32	0	0	0.0010	0.0010	1.0000
MamGypLTR1a LTR Gypsy	170.2	137	0	0	0.0010	0.0010	1.0000
UCON11 Unknown Unknown	184.7	45	0	0	0.0010	0.0010	1.0000
UCON27 Unknown Unknown	221	64	0	0	0.0010	0.0010	1.0000
LTR23-int LTR ERV1	81.6	11	0	0	0.0010	0.0010	1.0000
Eulor4 Unknown Unknown	209.8	24	0	0	0.0010	0.0010	1.0000
LTR29 LTR ERV1	164.5	34	0	0	0.0010	0.0010	1.0000
LTR86A2 LTR ERV1	154.7	79	0	0	0.0010	0.0010	1.0000
MER83B-int LTR ERV1	113.8	9	0	0	0.0010	0.0010	1.0000
RMER3D1 LTR ERV1	90	1	0	0	0.0010	0.0010	1.0000
RLTR43A LTR ERV1	92.7	3	0	0	0.0010	0.0010	1.0000
MER134 DNA? DNA?	131.7	53	0	0	0.0010	0.0010	1.0000
UCON22 Unknown Unknown	115.1	27	0	0	0.0010	0.0010	1.0000
MLT1E3-int LTR ERV1-MaLR	257.2	12	0	0	0.0010	0.0010	1.0000
MER127 DNA TcMar-Tigger	180.3	119	0	0	0.0010	0.0010	1.0000
L1PB4 LINE L1	335.7	144	0	0	0.0010	0.0010	1.0000
MER125 DNA DNA	123.8	111	0	0	0.0010	0.0010	1.0000
Ricksha_c DNA MULE-MuDR	173.7	26	0	0	0.0010	0.0010	1.0000
HY1 scRNA scRNA	90.6	38	0	0	0.0010	0.0010	1.0000
MER34-int LTR ERV1	263.1	66	0	0	0.0010	0.0010	1.0000
LTR36 LTR ERV1	99.2	5	0	0	0.0010	0.0010	1.0000
Ricksha_0 DNA MULE-MuDR	197.4	29	0	0	0.0010	0.0010	1.0000
MER68C LTR ERV1	158.8	60	0	0	0.0010	0.0010	1.0000
MLT1E1-int LTR ERV1-MaLR	413.6	8	0	0	0.0010	0.0010	1.0000
Eulor1 DNA DNA	169.2	56	0	0	0.0010	0.0010	1.0000
Ricksha_a DNA MULE-MuDR	132	11	0	0	0.0010	0.0010	1.0000
LTR91 LTR ERV1	134.7	57	0	0	0.0010	0.0010	1.0000



## Appendix

UCON20 Unknown Unknown	208.6	37	0	0	0.0010	0.0010	1.0000
MLT1I-int LTR ERV1-MaLR	265	57	0	0	0.0010	0.0010	1.0000
Eulor9B DNA DNA	147.1	14	0	0	0.0010	0.0010	1.0000
MER92C LTR ERV1	191.7	72	0	0	0.0010	0.0010	1.0000
LTR73 LTR ERV1	236.2	137	0	0	0.0010	0.0010	1.0000
LTR52 LTR ERV1	198.1	230	0	0	0.0010	0.0010	1.0000
Charlie10a DNA hAT-Charlie	139.7	61	0	0	0.0010	0.0010	1.0000
HERVL74-int LTR ERV1	290.6	149	0	0	0.0010	0.0010	1.0000
HY4 scRNA scRNA	82.7	3	0	0	0.0010	0.0010	1.0000
LTR32 LTR ERV1	115.8	9	0	0	0.0010	0.0010	1.0000
LTR64 LTR ERV1	105.2	8	0	0	0.0010	0.0010	1.0000
MER65-int LTR ERV1	167.1	49	0	0	0.0010	0.0010	1.0000
MER110 LTR ERV1	195.8	147	0	0	0.0010	0.0010	1.0000
Eulor8 DNA TcMar?	200.2	110	0	0	0.0010	0.0010	1.0000
UCON21 DNA? DNA?	128.3	23	0	0	0.0010	0.0010	1.0000
MER4CL34 LTR ERV1	68.8	6	0	0	0.0010	0.0010	1.0000
LTR108e_Mam LTR ERV1	210	36	0	0	0.0010	0.0010	1.0000
RLTR46A2 LTR ERV1	290.7	58	0	0	0.0010	0.0010	1.0000
MER70-int LTR ERV1	282.4	56	0	0	0.0010	0.0010	1.0000
Helitron1Na_Mam RC Helitron	229.8	62	0	0	0.0010	0.0010	1.0000
MER92A LTR ERV1	209.5	37	0	0	0.0010	0.0010	1.0000
MamRep1151 LTR? LTR?	271.9	249	0	0	0.0010	0.0010	1.0000
MamGypLTR3 LTR Gypsy	155.1	105	0	0	0.0010	0.0010	1.0000
Eulor5B DNA? DNA?	114.3	37	0	0	0.0010	0.0010	1.0000
RLTR9A4 LTR ERV1	246.4	35	0	0	0.0010	0.0010	1.0000
MER76-int LTR ERV1	250.6	16	0	0	0.0010	0.0010	1.0000
RLTR43B LTR ERV1	126	1	0	0	0.0010	0.0010	1.0000
UCON18 Unknown Unknown	174.2	17	0	0	0.0010	0.0010	1.0000
Eulor6D DNA? DNA?	151.1	47	0	0	0.0010	0.0010	1.0000
MER89-int LTR ERV1	263.6	98	0	0	0.0010	0.0010	1.0000
L1M2a1 LINE L1	63	11	0	0	0.0010	0.0010	1.0000
MLT1A-int LTR ERV1-MaLR	286.3	123	0	0	0.0010	0.0010	1.0000
MER57E2 LTR ERV1	190.1	62	0	0	0.0010	0.0010	1.0000
Charlie20a DNA hAT-Charlie	223.7	161	0	0	0.0010	0.0010	1.0000
UCON24 Unknown Unknown	131.3	18	0	0	0.0010	0.0010	1.0000
LTR103_Mam LTR ERV1?	183.6	58	0	0	0.0010	0.0010	1.0000
Eulor7 DNA? DNA?	142.8	4	0	0	0.0010	0.0010	1.0000
Charlie17a DNA hAT-Charlie	136.1	63	0	0	0.0010	0.0010	1.0000
UCON8 Unknown Unknown	164	55	0	0	0.0010	0.0010	1.0000
L1M3a LINE L1	257.7	120	0	0	0.0010	0.0010	1.0000
Eulor2C DNA? DNA?	120.6	24	0	0	0.0010	0.0010	1.0000
MLT1E-int LTR ERV1-MaLR	266.2	9	0	0	0.0010	0.0010	1.0000
Eulor12 DNA? DNA?	117.8	35	0	0	0.0010	0.0010	1.0000
LTR75 LTR ERV1	211.9	94	0	0	0.0010	0.0010	1.0000
MER72B LTR ERV1	111.5	74	0	0	0.0010	0.0010	1.0000
MamRep1161 DNA TcMar	170.4	7	0	0	0.0010	0.0010	1.0000
MER101B LTR ERV1	133.3	21	0	0	0.0010	0.0010	1.0000
MamRep137 DNA TcMar?	115.5	4	0	0	0.0010	0.0010	1.0000
MLT1G3-int LTR ERV1-MaLR	366.2	58	0	0	0.0010	0.0010	1.0000
UCON9 Unknown Unknown	184.7	28	0	0	0.0010	0.0010	1.0000
Eulor6C DNA? DNA?	116.6	15	0	0	0.0010	0.0010	1.0000
MLT1O-int LTR ERV1-MaLR	174	2	0	0	0.0010	0.0010	1.0000
RLTR1E_MM LTR ERV1	223.7	86	0	0	0.0010	0.0010	1.0000
LTR86C LTR ERV1	204.2	75	0	0	0.0010	0.0010	1.0000
LTR103b_Mam LTR ERV1?	159.7	57	0	0	0.0010	0.0010	1.0000
MLT1N2-int LTR ERV1-MaLR	339.7	3	0	0	0.0010	0.0010	1.0000
RMER3D4 LTR ERV1	63	2	0	0	0.0010	0.0010	1.0000
UCON16 Unknown Unknown	136.7	23	0	0	0.0010	0.0010	1.0000
X7D_LINE LINE CR1	61.5	10	0	0	0.0010	0.0010	1.0000
MER51-int LTR ERV1	173.1	7	0	0	0.0010	0.0010	1.0000

## Appendix

Cheshire_Mars DNA hAT-Charlie	202.7	3	0	0	0.0010	0.0010	1.0000
Tigger1a_Mars DNA TcMar-Tigger	253	2	0	0	0.0010	0.0010	1.0000
MLT1J2-int LTR ERV1-MaLR	302.7	27	0	0	0.0010	0.0010	1.0000
ERV4_1-LTR_MM LTR ERV1	150.9	21	0	0	0.0010	0.0010	1.0000
MER96B DNA hAT	119.1	8	0	0	0.0010	0.0010	1.0000
LTR85b LTR Gypsy?	191.6	106	0	0	0.0010	0.0010	1.0000
MER66B LTR ERV1	81.7	7	0	0	0.0010	0.0010	1.0000
Ricksha_b DNA MULE-MuDR	262.3	6	0	0	0.0010	0.0010	1.0000
HUERS-P3b-int LTR ERV1	88.3	3	0	0	0.0010	0.0010	1.0000
Charlie11 DNA hAT-Charlie	205.5	52	0	0	0.0010	0.0010	1.0000
LTR82B LTR ERV1	210.8	290	0	0	0.0010	0.0010	1.0000
LTR59 LTR ERV1	88.5	8	0	0	0.0010	0.0010	1.0000
ORR1C2 LTR ERV1-MaLR	238.7	12473	404.421	402.999	0.0011	0.0011	0.9996
L1MED LINE L1	217	3245	263.128	262.25	0.0014	0.0014	0.9991
LTR16A1 LTR ERV1	198.8	654	33.843	33.683	0.0013	0.0013	0.9990
IAPEY5_I-int LTR ERV1	663	184	2.538	2.406	0.0010	0.0010	0.9989
RLTR10E LTR ERV1	298.2	227	25.382	25.263	0.0014	0.0014	0.9987
RLTR9C LTR ERV1	366.8	477	5.076	4.812	0.0010	0.0010	0.9985
L1Md_Gf LINE L1	2209	1080	42.303	38.495	0.0010	0.0010	0.9984
L1MA7 LINE L1	262	3250	50.764	49.322	0.0011	0.0011	0.9984
MER63D DNA hAT-Blackjack	164.9	1034	13.537	13.233	0.0011	0.0011	0.9983
RLTR17C_MM LTR ERV1	352.2	218	2.538	2.406	0.0010	0.0010	0.9983
RLTR47 LTR ERV1	330.6	209	2.538	2.406	0.0010	0.0010	0.9982
MLTR31A_MM LTR ERV1	387.7	352	5.076	4.812	0.0010	0.0010	0.9981
MMAR1 DNA TcMar-Mariner	344.2	309	27.92	27.669	0.0013	0.0013	0.9981
Zaphod DNA hAT-Tip100	267.1	728	7.615	7.218	0.0010	0.0010	0.9980
L1MC1 LINE L1	299.9	5571	200.518	196.086	0.0011	0.0011	0.9976
RLTR13D5 LTR ERV1	758.4	745	10.999	9.624	0.0010	0.0010	0.9976
MER44D DNA TcMar-Tigger	176.7	940	16.075	15.639	0.0011	0.0011	0.9976
MER70A LTR ERV1	282.8	184	2.538	2.406	0.0010	0.0010	0.9976
RLTR48A LTR ERV1	422.9	635	21.152	20.451	0.0011	0.0011	0.9976
RMER12 LTR ERV1	297.9	3319	138.755	135.937	0.0011	0.0011	0.9975
L1Md_A LINE L1	1939.8	16844	1047.434	955.167	0.0010	0.0010	0.9973
MER21C LTR ERV1	267.8	2261	126.064	123.907	0.0012	0.0012	0.9971
RLTR13C3 LTR ERV1	755.4	538	5.076	3.609	0.0010	0.0010	0.9964
RLTR44D LTR ERV1	343.4	100	2.538	2.406	0.0011	0.0011	0.9964
MER63C DNA hAT-Blackjack	329.2	414	10.153	9.624	0.0011	0.0011	0.9964
MLT2C2 LTR ERV1	197.3	651	1.692	1.203	0.0010	0.0010	0.9962
Lx2 LINE L1	562.4	18521	472.953	430.667	0.0010	0.0010	0.9961
L1MEg LINE L1	219.3	3797	560.098	554.574	0.0017	0.0017	0.9960
RLTR10 LTR ERV1	377.9	2930	201.365	196.086	0.0012	0.0012	0.9960
MTB_MM LTR ERV1-MaLR	312.8	4963	225.054	217.74	0.0011	0.0011	0.9959
ORR1C1 LTR ERV1-MaLR	264.4	7423	269.05	259.844	0.0011	0.0011	0.9959
B1_Mur1 SINE Alu	122.9	38058	2493.366	2463.705	0.0015	0.0015	0.9959
Lx2B LINE L1	620.2	6191	115.912	98.644	0.0010	0.0010	0.9956
MER45B DNA hAT-Tip100	237.4	688	6.769	6.015	0.0010	0.0010	0.9956
L1_Mur1 LINE L1	581.8	7793	53.302	32.48	0.0010	0.0010	0.9955
MER103C DNA hAT-Charlie	142.3	1072	65.993	64.961	0.0014	0.0014	0.9953
L1M2 LINE L1	293.2	18488	336.736	309.166	0.0011	0.0011	0.9952
MT2B1 LTR ERV1	270.8	7248	160.753	150.373	0.0011	0.0011	0.9951
L1Md_F2 LINE L1	1140.7	64898	1973.034	1598.762	0.0010	0.0010	0.9951
MamGypLTR1b LTR Gypsy	166.8	141	2.538	2.406	0.0011	0.0011	0.9949
Lx3_Mus LINE L1	694.3	12651	250.437	204.507	0.0010	0.0010	0.9949
LTR33 LTR ERV1	191.5	1466	13.537	12.03	0.0010	0.0010	0.9949
L1Md_F3 LINE L1	920.5	16146	478.875	396.984	0.0010	0.0010	0.9947
MLTR18D_MM LTR ERV1	429.1	620	21.998	20.451	0.0011	0.0011	0.9946
RLTR25A LTR ERV1	295.4	1781	86.299	83.006	0.0012	0.0012	0.9946
MLTR31FA_MM LTR ERV1	417	525	16.921	15.639	0.0011	0.0011	0.9946
L1MB5 LINE L1	260.9	3766	192.904	186.462	0.0012	0.0012	0.9945
MTA_MM LTR ERV1-MaLR	370.3	15623	1053.356	1014.113	0.0012	0.0012	0.9943

## Appendix

LTR82A LTR ERV L	237.8	341	1.692	1.203	0.0010	0.0010	0.9941
RLTR13B1 LTR ERV K	723.4	1029	34.689	30.075	0.0010	0.0010	0.9941
RLTR25B LTR ERV K	237.9	3674	186.981	180.447	0.0012	0.0012	0.9938
L1_Mur3 LINE L1	465	27399	434.034	352.473	0.0010	0.0010	0.9938
LTR88c LTR Gypsy?	157.5	118	2.538	2.406	0.0011	0.0011	0.9938
RLTR45-int LTR ERV K	696.6	1317	306.277	298.339	0.0013	0.0013	0.9935
L1MC4 LINE L1	228.8	5589	162.445	152.779	0.0011	0.0011	0.9933
MER63B DNA hAT-Blackjack	150.6	701	6.769	6.015	0.0011	0.0011	0.9933
ORR1A0 LTR ERV L-MaLR	316.5	2133	137.909	132.328	0.0012	0.0012	0.9931
LTR16A2 LTR ERV L	203.8	416	4.23	3.609	0.0010	0.0010	0.9930
LTR16C LTR ERV L	199.4	1352	15.229	13.233	0.0011	0.0010	0.9930
IAPEY3_LTR LTR ERV K	302	553	8.461	7.218	0.0011	0.0010	0.9929
RMER13B LTR ERV K	557.5	2134	120.142	110.674	0.0011	0.0011	0.9928
L1_Mus4 LINE L1	738.3	11880	221.67	155.185	0.0010	0.0010	0.9926
RMER6C LTR ERV K	396.5	4887	105.759	90.224	0.0011	0.0010	0.9924
MLTR73 LTR ERV 1	170.6	956	19.46	18.045	0.0011	0.0011	0.9922
RCHARR1 DNA hAT-Charlie	308.8	6811	297.816	279.092	0.0011	0.0011	0.9922
Tigger12 DNA TcMar-Tigger	170.2	155	49.918	49.322	0.0029	0.0029	0.9922
Lx6 LINE L1	503.8	18164	459.416	383.751	0.0011	0.0010	0.9921
L1VL4 LINE L1	1032.4	6091	172.598	121.501	0.0010	0.0010	0.9921
MURVY-LTR LTR ERV 1	449.3	1910	39.765	32.48	0.0010	0.0010	0.9919
RLTR24B_MM LTR ERV 1	192.2	1018	18.614	16.842	0.0011	0.0011	0.9917
RLTR13A1 LTR ERV K	613.1	205	5.922	4.812	0.0010	0.0010	0.9916
Lx5c LINE L1	460.1	16422	405.267	338.038	0.0011	0.0010	0.9916
RMER2 LTR ERV 1	356.1	2358	53.302	45.713	0.0011	0.0011	0.9915
L1_Mur2 LINE L1	537.7	20769	463.646	364.503	0.0010	0.0010	0.9915
L1MA10 LINE L1	207.3	866	7.615	6.015	0.0010	0.0010	0.9915
RLTR22_Mur LTR ERV K	445	2503	72.762	62.555	0.0011	0.0011	0.9914
Lx3A LINE L1	660.8	9223	168.368	114.283	0.0010	0.0010	0.9914
L1MA5 LINE L1	258.8	4952	108.297	96.238	0.0011	0.0011	0.9913
L1Md_T LINE L1	2054.8	23688	1406.167	970.806	0.0010	0.0010	0.9913
LTR53 LTR ERV L	240.4	226	1.692	1.203	0.0010	0.0010	0.9913
L1MEc LINE L1	268.5	5378	124.372	110.674	0.0011	0.0011	0.9913
L1_Mus3 LINE L1	757	48489	1099.044	766.299	0.0010	0.0010	0.9912
RLTR9A3B LTR ERV K	271.5	346	0.846	0	0.0010	0.0010	0.9911
HERV16-int LTR ERV L	234.9	277	4.23	3.609	0.0011	0.0011	0.9910
L1_Mm LINE L1	643.3	5429	115.912	83.006	0.0010	0.0010	0.9909
ORR1D-int LTR ERV L-MaLR	457.5	878	10.999	7.218	0.0010	0.0010	0.9908
MLTR14 LTR ERV 1	251.3	6798	189.52	172.026	0.0011	0.0011	0.9908
Lx LINE L1	658.5	24834	659.934	502.846	0.0010	0.0010	0.9908
L1MA4 LINE L1	292.4	7467	219.978	197.289	0.0011	0.0011	0.9906
MER68 LTR ERV L	265	637	57.533	55.337	0.0013	0.0013	0.9903
RLTR17D_Mm LTR ERV K	596.4	1225	55.841	48.119	0.0011	0.0011	0.9902
LTR16E2 LTR ERV L	198.3	423	0.846	0	0.0010	0.0010	0.9900
Lx3C LINE L1	601.9	6475	170.906	129.922	0.0010	0.0010	0.9899
Lx4B LINE L1	533.1	5680	274.126	240.596	0.0011	0.0011	0.9898
ERV B7_2B-LTR_MM LTR ERV K	328.8	710	27.92	25.263	0.0011	0.0011	0.9898
L1ME4b LINE L1	188.3	1818	111.681	107.065	0.0013	0.0013	0.9898
Lx9 LINE L1	323	49160	937.445	762.69	0.0011	0.0010	0.9896
Lx2B2 LINE L1	576	11012	174.29	105.862	0.0010	0.0010	0.9895
MER33 DNA hAT-Charlie	171.1	2719	88.837	83.006	0.0012	0.0012	0.9895
LTR16E1 LTR ERV L	190.5	514	5.922	4.812	0.0011	0.0010	0.9893
MTEa LTR ERV L-MaLR	221.1	20061	518.641	464.351	0.0011	0.0011	0.9890
LTR67B LTR ERV L	168.2	632	19.46	18.045	0.0012	0.0012	0.9887
RLTR42 LTR ERV K	157	266	1.692	1.203	0.0010	0.0010	0.9887
RMER19A LTR ERV K	484.8	2095	104.067	91.427	0.0011	0.0011	0.9887
L1MB8 LINE L1	263.3	6000	335.89	313.978	0.0012	0.0012	0.9886
ORR1A2-int LTR ERV L-MaLR	828.9	2132	199.672	176.838	0.0011	0.0011	0.9884
ORR1D1 LTR ERV L-MaLR	235.8	21507	931.522	861.335	0.0012	0.0012	0.9883
L3 LINE CR1	129.6	9732	526.255	505.252	0.0014	0.0014	0.9883

## Appendix

RMER19B LTR ERVK	543.5	5640	479.721	437.885	0.0012	0.0011	0.9882
Lx3B LINE L1	570.6	5993	243.668	199.695	0.0011	0.0011	0.9880
RMER12B LTR ERVK	414.7	1011	146.37	139.546	0.0013	0.0013	0.9879
MT2_Mm LTR ERV	450.9	2667	135.371	119.095	0.0011	0.0011	0.9878
MTE2b-int LTR ERV-MaLR	430	729	35.535	31.278	0.0011	0.0011	0.9878
Lx7 LINE L1	404.1	40190	939.983	725.398	0.0011	0.0010	0.9875
Lx5 LINE L1	585.9	16339	1370.632	1233.056	0.0011	0.0011	0.9874
MTC-int LTR ERV-MaLR	424.8	3520	263.974	241.799	0.0012	0.0012	0.9874
Lx10 LINE L1	291	10010	170.06	131.125	0.0011	0.0010	0.9874
RMER1C Other Other	312.3	6347	151.446	123.907	0.0011	0.0011	0.9871
L1MA9 LINE L1	268.2	4587	51.61	34.886	0.0010	0.0010	0.9870
RMER20B LTR ERVK	310.8	1949	28.766	20.451	0.0010	0.0010	0.9869
Lx4A LINE L1	568.7	10564	379.039	294.73	0.0011	0.0010	0.9868
PB1 SINE Alu	97.1	11784	758.078	732.616	0.0017	0.0016	0.9866
RLTR13D4 LTR ERVK	595.2	205	1.692	0	0.0010	0.0010	0.9863
L1ME3A LINE L1	237.7	2086	155.677	146.764	0.0013	0.0013	0.9863
RodERV21-int LTR ERV1	577.4	1856	68.532	52.931	0.0011	0.0010	0.9863
B1_Mm SINE Alu	131.1	42910	3303.055	3180.682	0.0016	0.0016	0.9863
ORR1E-int LTR ERV-MaLR	461.5	1446	31.305	21.654	0.0010	0.0010	0.9862
Tigger5b DNA TcMar-Tigger	142.5	2473	43.996	38.495	0.0011	0.0011	0.9861
Tigger13a DNA TcMar-Tigger	229	645	11.845	9.624	0.0011	0.0011	0.9861
MTE2a LTR ERV-MaLR	258.8	15148	580.404	517.282	0.0011	0.0011	0.9860
L1MC3 LINE L1	291.5	5288	118.45	95.036	0.0011	0.0011	0.9859
MER89 LTR ERV1	245.1	614	3.384	1.203	0.0010	0.0010	0.9858
RLTR11A2 LTR ERVK	369.8	2676	309.661	291.121	0.0013	0.0013	0.9857
IAPLTR2b LTR ERVK	309.9	1017	44.842	39.698	0.0011	0.0011	0.9857
RMER21B LTR ERV1	230.2	1252	34.689	30.075	0.0011	0.0011	0.9857
MER104 DNA TcMar-Tc2	131.2	295	4.23	3.609	0.0011	0.0011	0.9855
Lx8b LINE L1	347.6	24166	707.314	575.025	0.0011	0.0011	0.9855
MER34A1 LTR ERV1	201.1	541	21.152	19.248	0.0012	0.0012	0.9853
Lx5b LINE L1	644.7	7975	251.283	172.026	0.0010	0.0010	0.9853
L1MEf LINE L1	256	3127	58.379	45.713	0.0011	0.0011	0.9853
RLTR13C2 LTR ERVK	596.7	844	38.073	30.075	0.0011	0.0011	0.9852
RLTR17 LTR ERVK	454.7	1758	126.91	113.08	0.0012	0.0011	0.9851
RMER15-int LTR ERV	365.3	3100	236.053	215.334	0.0012	0.0012	0.9849
IAPEz-int LTR ERVK	1622.6	7319	3006.93	2781.292	0.0013	0.0012	0.9848
L1MA8 LINE L1	271.1	3882	83.761	66.164	0.0011	0.0011	0.9845
RLTR20B3 LTR ERVK	309.5	974	41.457	36.089	0.0011	0.0011	0.9843
RLTR1D LTR ERV1	478.5	305	41.457	38.495	0.0013	0.0013	0.9842
RMER17C2 LTR ERVK	207.2	3413	54.148	42.104	0.0011	0.0011	0.9842
IAPLTR3 LTR ERVK	330.8	959	7.615	2.406	0.0010	0.0010	0.9840
MIRc SINE MIR	104.1	23213	1254.721	1195.763	0.0015	0.0015	0.9839
ORR1F LTR ERV-MaLR	209.1	23086	555.021	467.96	0.0011	0.0011	0.9838
MLT1C-int LTR ERV-MaLR	279.5	122	4.23	3.609	0.0011	0.0011	0.9838
MTB LTR ERV-MaLR	333.6	6419	416.266	374.127	0.0012	0.0012	0.9835
LTR41B LTR ERV	192.4	290	3.384	2.406	0.0011	0.0010	0.9835
MMERVK10D3_LTR LTR ERVK	320.6	174	3.384	2.406	0.0011	0.0010	0.9835
MER46C DNA TcMar-Tigger	162.3	788	14.383	12.03	0.0011	0.0011	0.9835
RLTR20A2B_MM LTR ERVK	299	438	16.921	14.436	0.0011	0.0011	0.9832
MurERV4_19-int LTR ERVK	627.3	969	28.766	18.045	0.0010	0.0010	0.9832
RMER17B LTR ERVK	580.9	2898	221.67	188.868	0.0011	0.0011	0.9828
LTR41 LTR ERV	208.8	506	17.767	15.639	0.0012	0.0011	0.9828
MLT1L LTR ERV-MaLR	178.3	1535	54.995	49.322	0.0012	0.0012	0.9827
ORR1A4 LTR ERV-MaLR	270.1	7791	856.222	804.794	0.0014	0.0014	0.9826
B1_Mur3 SINE Alu	124.1	24632	1858.814	1773.194	0.0016	0.0016	0.9826
ORR1C2-int LTR ERV-MaLR	534.5	914	62.609	52.931	0.0011	0.0011	0.9824
LTR40a LTR ERV	234.8	709	4.23	1.203	0.0010	0.0010	0.9823
IAPEY2_LTR LTR ERVK	322.7	1062	13.537	7.218	0.0010	0.0010	0.9823
ORR1A3 LTR ERV-MaLR	259.9	4381	301.201	275.483	0.0013	0.0012	0.9821
RLTR12E LTR ERVK	361.3	615	45.688	40.901	0.0012	0.0012	0.9821

## Appendix

L1M2c LINE L1	304	321	4.23	2.406	0.0010	0.0010	0.9821
B2_Mm1a SINE B2	168.5	18346	6537.578	6364.973	0.0031	0.0031	0.9821
ERVb7_4-LTR_MM LTR ERVk	335.9	166	5.922	4.812	0.0011	0.0011	0.9820
RLTR16C_MM LTR ERVk	377.4	360	13.537	10.827	0.0011	0.0011	0.9819
L1M3 LINE L1	241.7	7043	136.217	102.253	0.0011	0.0011	0.9815
MER2 DNA TcMar-Tigger	184.6	10246	302.893	262.25	0.0012	0.0011	0.9815
LTR16A LTR ERVl	209.9	1472	16.921	10.827	0.0011	0.0010	0.9813
LTR78 LTR ERV1	200.2	583	5.922	3.609	0.0011	0.0010	0.9811
MER74B LTR ERVl	237.9	324	7.615	6.015	0.0011	0.0011	0.9811
MER5A DNA hAT-Charlie	109.7	8982	222.516	199.695	0.0012	0.0012	0.9811
MLT1F2 LTR ERVl-MaLR	243.5	2297	76.992	64.961	0.0011	0.0011	0.9811
ERVb3_1-LTR_MM LTR ERVk	276.3	158	0.846	0	0.0010	0.0010	0.9810
MER77 LTR ERVl	246.7	570	21.152	18.045	0.0012	0.0011	0.9808
L1MC5a LINE L1	195.4	2588	76.146	64.961	0.0012	0.0011	0.9808
B1_Mus2 SINE Alu	128.1	70891	5892.027	5599.877	0.0016	0.0016	0.9805
ORR1A1 LTR ERVl-MaLR	284.5	4414	219.132	190.071	0.0012	0.0012	0.9803
L1ME2z LINE L1	233.6	1533	35.535	27.669	0.0011	0.0011	0.9800
ORR1B2 LTR ERVl-MaLR	259	8528	357.041	305.557	0.0012	0.0011	0.9799
L1MA5A LINE L1	335.7	1231	46.534	37.292	0.0011	0.0011	0.9799
B1_Mus1 SINE Alu	133.3	95081	7075.678	6676.546	0.0016	0.0015	0.9798
RMER6A LTR ERVk	449.8	3983	369.732	326.008	0.0012	0.0012	0.9798
LTR37A LTR ERV1	189.9	549	3.384	1.203	0.0010	0.0010	0.9797
IAPEy-int LTR ERVk	3099.7	755	69.378	20.451	0.0010	0.0010	0.9797
MER57C2 LTR ERV1	193.6	415	1.692	0	0.0010	0.0010	0.9794
RLTR51A_Mm LTR ERVk?	132.7	3523	51.61	40.901	0.0011	0.0011	0.9794
MERVL-int LTR ERVl	1660.9	2600	401.037	303.151	0.0011	0.0011	0.9793
Tigger17c DNA TcMar-Tigger	201.3	649	10.153	7.218	0.0011	0.0011	0.9792
RLTR1D2_MM LTR ERV1	492.8	192	11.845	9.624	0.0011	0.0011	0.9791
RLTR20B2 LTR ERVk	301.8	880	49.918	43.307	0.0012	0.0012	0.9790
RLTR21 LTR ERVk	355.8	3816	171.752	139.546	0.0011	0.0011	0.9789
MTE-int LTR ERVl-MaLR	348.8	4958	520.333	472.772	0.0013	0.0013	0.9789
MER57E1 LTR ERV1	186.2	210	0.846	0	0.0010	0.0010	0.9788
Kanga1a DNA TcMar-Tc2	204.5	191	0.846	0	0.0010	0.0010	0.9788
MLT1E3 LTR ERVl-MaLR	227.5	1020	29.612	24.06	0.0011	0.0011	0.9788
RLTR44B LTR ERVk	307.9	275	6.769	4.812	0.0011	0.0011	0.9786
L2 LINE L2	196.1	17614	632.86	544.95	0.0012	0.0012	0.9785
RLTR13B3 LTR ERVk	816.6	265	8.461	3.609	0.0010	0.0010	0.9784
RLTR41A2 LTR ERV1	353.1	440	5.922	2.406	0.0010	0.0010	0.9782
RMER15 LTR ERVl	219.7	18077	395.96	300.745	0.0011	0.0011	0.9782
MARNA DNA TcMar-Mariner	179.4	562	32.997	30.075	0.0013	0.0013	0.9782
RLTR44E LTR ERVk	338.3	122	3.384	2.406	0.0011	0.0011	0.9781
ORR1B2-int LTR ERVl-MaLR	564.9	857	62.609	50.525	0.0011	0.0011	0.9779
RSINE1 SINE B4	124.1	11580 6	4907.202	4479.902	0.0013	0.0013	0.9778
L1MC LINE L1	233.3	5426	202.211	169.62	0.0012	0.0011	0.9778
BLACKJACK DNA hAT- Blackjack	302.3	716	13.537	8.421	0.0011	0.0010	0.9778
RLTR24 LTR ERV1	326.7	1240	36.381	26.466	0.0011	0.0011	0.9775
MLT2F LTR ERVl	219.6	574	55.841	51.728	0.0014	0.0014	0.9774
L1M3e LINE L1	279.8	652	31.305	26.466	0.0012	0.0011	0.9774
RLTR26 LTR ERVk	341.6	408	11.845	8.421	0.0011	0.0011	0.9774
L1M4 LINE L1	242.6	13935	493.258	405.405	0.0011	0.0011	0.9773
MTC LTR ERVl-MaLR	274	27489	1220.032	1021.331	0.0012	0.0011	0.9773
RMER4B LTR ERVk	303.7	5158	368.886	324.805	0.0012	0.0012	0.9772
L1M5 LINE L1	213.7	16596	432.341	340.444	0.0011	0.0011	0.9769
RLTR13D1 LTR ERVk	573.1	255	5.922	2.406	0.0010	0.0010	0.9769
Lx8 LINE L1	334.1	49141	1907.04	1480.87	0.0011	0.0011	0.9767
MER68-int LTR ERVl	251	77	1.692	1.203	0.0011	0.0011	0.9767
RMER17A2 LTR ERVk	459.5	2598	231.823	198.492	0.0012	0.0012	0.9766
RLTR14_RN LTR ERV1	256.5	1148	30.458	22.857	0.0011	0.0011	0.9766
RMER20A LTR ERVk	292.8	1559	54.148	42.104	0.0011	0.0011	0.9764

## Appendix

MLT1G3 LTR ERV-L-MaLR	221	1075	27.92	21.654	0.0011	0.0011	0.9764
RLTR11C_MM LTR ERV-K	357.4	858	43.15	34.886	0.0011	0.0011	0.9764
RMER19C LTR ERV-K	408.6	2729	54.148	26.466	0.0010	0.0010	0.9763
IAP1-MM_I-int LTR ERV-K	827	479	21.998	12.03	0.0011	0.0010	0.9762
RLTR28 LTR ERV-L	372.4	1949	258.051	234.581	0.0014	0.0013	0.9761
MTE2b LTR ERV-L-MaLR	232.4	14367	432.341	341.647	0.0011	0.0011	0.9760
Tigger20a DNA TcMar-Tigger	196.1	188	3.384	2.406	0.0011	0.0011	0.9757
MMERV-K9C_I-int LTR ERV-K	1074.5	1807	332.505	276.686	0.0012	0.0011	0.9755
RLTR20C LTR ERV-K	227.8	263	10.153	8.421	0.0012	0.0011	0.9753
ORR1A3-int LTR ERV-L-MaLR	320.4	2435	82.915	61.352	0.0011	0.0011	0.9750
ORR1B1 LTR ERV-L-MaLR	296.4	17313	654.858	510.064	0.0011	0.0011	0.9750
RMER10B LTR ERV-L	248	3002	203.057	179.244	0.0013	0.0012	0.9749
B1F1 SINE Alu	109	19678	1257.259	1171.704	0.0016	0.0015	0.9749
ERV-L-B4-int LTR ERV-L	323.1	1384	33.843	21.654	0.0011	0.0010	0.9747
IAP1-MM_LTR LTR ERV-K	272.6	420	4.23	1.203	0.0010	0.0010	0.9745
Tigger14a DNA TcMar-Tigger	143.4	225	0.846	0	0.0010	0.0010	0.9744
ORR1A2 LTR ERV-L-MaLR	279.3	14843	725.081	599.085	0.0012	0.0011	0.9741
RLTR13C1 LTR ERV-K	712.4	309	19.46	13.233	0.0011	0.0011	0.9740
ERV-B2_1A-I_MM-int LTR ERV-K	574.7	289	11.845	7.218	0.0011	0.0010	0.9740
MER63A DNA hAT-Blackjack	156.5	635	7.615	4.812	0.0011	0.0010	0.9738
RLTR13G LTR ERV-K	347.6	877	62.609	52.931	0.0012	0.0012	0.9737
L1MA6 LINE L1	275.5	6985	143.832	89.021	0.0011	0.0010	0.9735
B4A SINE B4	181.7	10707 3	5533.294	4869.668	0.0013	0.0013	0.9734
RLTR17B_Mm LTR ERV-K	136.5	27369	566.02	451.118	0.0012	0.0011	0.9733
RMER16 LTR ERV-K	348.8	471	46.534	40.901	0.0013	0.0012	0.9733
CYRA11_Mm Unknown Y-chromosome	142.8	2869	21.152	9.624	0.0011	0.0010	0.9732
L1MD2 LINE L1	261.6	3227	164.137	137.14	0.0012	0.0012	0.9732
MLT2C1 LTR ERV-L	203.1	872	13.537	8.421	0.0011	0.0010	0.9732
L1ME1 LINE L1	279.3	5261	263.128	216.537	0.0012	0.0011	0.9731
RMER1B Other Other	461.4	7511	606.632	496.831	0.0012	0.0011	0.9730
MamRep4096 DNA hAT?	169.3	180	0.846	0	0.0010	0.0010	0.9730
MTD LTR ERV-L-MaLR	259.8	51544	2064.409	1645.678	0.0012	0.0011	0.9729
L1MB7 LINE L1	269.6	6956	390.038	328.414	0.0012	0.0012	0.9728
Charlie12 DNA hAT-Charlie	196	902	16.075	10.827	0.0011	0.0011	0.9728
LTRIS2 LTR ERV-1	458.5	1423	118.45	97.441	0.0012	0.0011	0.9727
L1M4b LINE L1	301.4	2190	216.594	192.477	0.0013	0.0013	0.9725
L1ME2 LINE L1	269.3	2633	76.992	55.337	0.0011	0.0011	0.9725
ORR1B1-int LTR ERV-L-MaLR	467.8	2965	227.593	182.853	0.0012	0.0011	0.9723
L2b LINE L2	123.7	13720	478.029	417.434	0.0013	0.0012	0.9721
ORR1C1-int LTR ERV-L-MaLR	575.8	783	41.457	27.669	0.0011	0.0011	0.9720
RMER13A1 LTR ERV-K	483.2	1082	76.992	60.149	0.0011	0.0011	0.9719
MLT1C LTR ERV-L-MaLR	216.5	7791	297.816	241.799	0.0012	0.0011	0.9718
MTEb LTR ERV-L-MaLR	208.3	10811	344.35	270.671	0.0012	0.0011	0.9716
L1ME3B LINE L1	225.5	735	32.151	26.466	0.0012	0.0012	0.9713
LTR16B2 LTR ERV-L	221.6	384	2.538	0	0.0010	0.0010	0.9710
L1MD1 LINE L1	294	2268	125.218	102.253	0.0012	0.0012	0.9710
RMER6B LTR ERV-K	297.4	1360	77.838	63.758	0.0012	0.0012	0.9708
L1MD LINE L1	223.3	6361	155.677	109.471	0.0011	0.0011	0.9707
MLT1H2 LTR ERV-L-MaLR	165.3	1034	10.153	4.812	0.0011	0.0010	0.9705
MLT1H LTR ERV-L-MaLR	198.3	2322	33.843	19.248	0.0011	0.0010	0.9705
MER67C LTR ERV-1	246.3	620	5.922	1.203	0.0010	0.0010	0.9703
PB1D9 SINE Alu	96.9	28939	2270.004	2118.45	0.0018	0.0018	0.9701
MLT1F-int LTR ERV-L-MaLR	249.9	417	14.383	10.827	0.0011	0.0011	0.9700
RLTR10B LTR ERV-K	201.6	477	22.844	19.248	0.0012	0.0012	0.9698
MLT1D LTR ERV-L-MaLR	219.2	8368	246.206	182.853	0.0011	0.0011	0.9695
Zaphod3 DNA hAT-Tip100	217.5	454	9.307	6.015	0.0011	0.0011	0.9695
RMER10A LTR ERV-L	272.1	4464	239.438	194.883	0.0012	0.0012	0.9694
RLTR19B LTR ERV-K	382	224	10.153	7.218	0.0011	0.0011	0.9693
MERV-L_2A-int LTR ERV-L	528	3899	309.661	236.987	0.0012	0.0011	0.9693



## Appendix

LTR80A LTR ERV L	236.4	112	0.846	0	0.0010	0.0010	0.9690
L1MD3 LINE L1	212.4	2480	82.915	63.758	0.0012	0.0011	0.9686
MER45C DNA hAT-Tip100	185.1	290	4.23	2.406	0.0011	0.0010	0.9685
RMER6BA LTR ERV K	234.3	957	32.151	24.06	0.0011	0.0011	0.9684
MER21-int LTR ERV L	246.2	105	0.846	0	0.0010	0.0010	0.9683
RLTR26_Mus LTR ERV K	480.1	1091	37.227	19.248	0.0011	0.0010	0.9680
LTR33C LTR ERV L	197.9	232	33.843	31.278	0.0017	0.0017	0.9678
LSU-rRNA_Hsa rRNA rRNA	146.1	481	4559.468	4410.129	0.0659	0.0638	0.9677
L1ME3Cz LINE L1	197.5	1051	14.383	7.218	0.0011	0.0010	0.9677
ORR1F-int LTR ERV L-MaLR	394.2	395	15.229	9.624	0.0011	0.0011	0.9672
RMER19B2 LTR ERV K	470.6	2296	164.137	122.704	0.0012	0.0011	0.9667
MER110A LTR ERV I	189.6	196	2.538	1.203	0.0011	0.0010	0.9664
RLTR12B LTR ERV K	349.6	909	58.379	45.713	0.0012	0.0011	0.9663
RLTR20B4_MM LTR ERV K	314	416	21.998	16.842	0.0012	0.0011	0.9662
MLT1J1 LTR ERV L-MaLR	157.5	1135	23.69	16.842	0.0011	0.0011	0.9662
MT2A LTR ERV L	255.2	10140	351.119	251.423	0.0011	0.0011	0.9661
MIR SINE MIR	127.4	38978	966.211	763.893	0.0012	0.0012	0.9659
ORR1D2-int LTR ERV L-MaLR	421.7	1361	131.141	107.065	0.0012	0.0012	0.9659
RLTR31C_MM LTR ERV K	302.9	726	65.147	55.337	0.0013	0.0013	0.9656
MTEa-int LTR ERV L-MaLR	463.6	1368	78.684	54.134	0.0011	0.0011	0.9656
RLTR13D LTR ERV K	521.9	45	0.846	0	0.0010	0.0010	0.9652
MLT1E1A LTR ERV L-MaLR	221.9	1154	30.458	20.451	0.0011	0.0011	0.9651
URR1A DNA hAT-Charlie	184.5	17621	840.147	696.526	0.0013	0.0012	0.9649
LTR105_Mam LTR ERV L	187.5	124	0.846	0	0.0010	0.0010	0.9649
Charlie25 DNA hAT-Charlie	271	301	4.23	1.203	0.0011	0.0010	0.9647
Arthur1B DNA hAT-Tip100	213.9	597	5.922	1.203	0.0010	0.0010	0.9647
MLTR11B LTR ERV K	289.6	2985	101.528	67.367	0.0011	0.0011	0.9646
LTR37-int LTR ERV I	253.3	239	8.461	6.015	0.0011	0.0011	0.9646
MIRb SINE MIR	128.4	42160	1455.239	1211.402	0.0013	0.0012	0.9645
B4 SINE B4	167.7	64020	2760.724	2279.649	0.0013	0.0012	0.9644
ORR1A0-int LTR ERV L-MaLR	1238.1	456	97.298	73.382	0.0012	0.0011	0.9639
L2a LINE L2	167.1	24277	988.209	804.794	0.0012	0.0012	0.9636
L1ME3 LINE L1	237.1	1527	47.38	32.48	0.0011	0.0011	0.9636
MLT1E2 LTR ERV L-MaLR	231.2	1821	82.069	63.758	0.0012	0.0012	0.9636
L1MB2 LINE L1	259.4	2927	254.667	217.74	0.0013	0.0013	0.9636
RMER5 LTR ERV I	280.2	7418	241.13	156.388	0.0011	0.0011	0.9635
L1M7 LINE L1	214.4	207	1.692	0	0.0010	0.0010	0.9633
UCON2 Unknown Unknown	141.3	82	1.692	1.203	0.0011	0.0011	0.9632
RMER13A LTR ERV K	518.4	1010	54.995	33.683	0.0011	0.0011	0.9632
RMER3D-int LTR ERV K	449.4	2711	549.945	484.801	0.0015	0.0014	0.9632
B1_Mur4 SINE Alu	122.3	41031	2681.194	2396.339	0.0015	0.0015	0.9630
MYSERV16_I-int LTR ERV K	568.4	2516	91.375	34.886	0.0011	0.0010	0.9629
RLTR13D3A LTR ERV K	553.6	162	8.461	4.812	0.0011	0.0011	0.9628
B3A SINE B2	149.4	92925	3842.001	3181.885	0.0013	0.0012	0.9628
Tigger1 DNA TcMar-Tigger	345.5	1025	57.533	42.104	0.0012	0.0011	0.9625
ERVB3_1-I_MM-int LTR ERV K	821.9	134	9.307	4.812	0.0011	0.0010	0.9624
LTR10_RN LTR ERV K	250.2	546	56.687	49.322	0.0014	0.0014	0.9619
RLTR10-int LTR ERV K	738.9	2617	241.13	157.591	0.0011	0.0011	0.9616
MLT1F LTR ERV L-MaLR	238.6	1561	28.766	13.233	0.0011	0.0010	0.9613
MamRep434 DNA TcMar-Tigger	185	518	38.919	33.683	0.0014	0.0014	0.9611
RLTR46A LTR ERV K	293.5	71	0.846	0	0.0010	0.0010	0.9610
RLTR31B2 LTR ERV K	292.2	1192	61.763	45.713	0.0012	0.0011	0.9609
L1M4a1 LINE L1	244.2	607	16.075	9.624	0.0011	0.0011	0.9607
MER124 DNA? DNA?	164	252	1.692	0	0.0010	0.0010	0.9607
RMER13A2 LTR ERV K	460.1	1473	109.143	78.194	0.0012	0.0011	0.9607
BGLII LTR ERV K	348.7	2269	610.016	554.574	0.0018	0.0017	0.9604
Ricksha_c DNA MULE-MuDR	153.6	405	5.076	2.406	0.0011	0.0010	0.9603
RLTR10D2 LTR ERV K	317.5	225	4.23	1.203	0.0011	0.0010	0.9600
B2_Mm1t SINE B2	164	23218	3859.768	3552.403	0.0020	0.0019	0.9599
ID_B1 SINE B4	130.7	11082	7677.233	6788.423	0.0015	0.0015	0.9599

## Appendix

		7					
MRLTR33 LTR ERVK	224.5	624	68.532	60.149	0.0015	0.0014	0.9598
MER113 DNA hAT-Charlie	189.5	643	75.3	67.367	0.0016	0.0016	0.9598
LTR33A_ LTR ERVK	196.8	474	12.691	8.421	0.0011	0.0011	0.9597
RMER30 DNA hAT-Charlie	122.5	3588	143.832	120.298	0.0013	0.0013	0.9597
BGLII_B2 LTR ERVK	271.2	359	342.658	324.805	0.0045	0.0043	0.9594
RLTR19D LTR ERVK	312.4	64	0.846	0	0.0010	0.0010	0.9594
MurERV4-int LTR ERVK	571.6	1525	284.279	236.987	0.0013	0.0013	0.9591
RLTR30C_MM LTR ERV1	318.9	124	1.692	0	0.0010	0.0010	0.9590
RLTR13D6 LTR ERVK	688.8	805	33.843	9.624	0.0011	0.0010	0.9588
MLT1I LTR ERV1-MaLR	169.3	1777	32.997	19.248	0.0011	0.0011	0.9588
MER31B LTR ERV1	217.5	558	17.767	12.03	0.0011	0.0011	0.9588
MLT1J2 LTR ERV1-MaLR	170.3	1110	16.921	8.421	0.0011	0.0010	0.9587
MTA_Mm-int LTR ERV1-MaLR	909.9	3027	667.549	525.703	0.0012	0.0012	0.9585
MER21B LTR ERV1	262.4	1618	125.218	102.253	0.0013	0.0012	0.9582
ORSL DNA hAT-Tip100	138.3	281	4.23	2.406	0.0011	0.0011	0.9577
RMER4A LTR ERVK	242.8	3287	346.888	298.339	0.0014	0.0014	0.9576
RMER17D2 LTR ERVK	414.1	1783	90.529	55.337	0.0011	0.0011	0.9575
RLTR9B2 LTR ERVK	324	178	5.076	2.406	0.0011	0.0010	0.9574
L1MB3 LINE L1	257.4	5343	249.59	180.447	0.0012	0.0011	0.9574
RLTR42-int LTR ERVK	268.6	2473	104.913	72.179	0.0012	0.0011	0.9574
MLT1B LTR ERV1-MaLR	203.4	9316	395.96	298.339	0.0012	0.0012	0.9574
LTR40A1 LTR ERV1	166.5	114	0.846	0	0.0010	0.0010	0.9573
RLTR20B3A_MM LTR ERVK	320.5	843	34.689	21.654	0.0011	0.0011	0.9572
B3 SINE B2	167.1	14842 7	8204.335	6790.829	0.0013	0.0013	0.9572
LTR88b LTR Gypsy?	160.5	178	2.538	1.203	0.0011	0.0010	0.9571
Tigger7 DNA TcMar-Tigger	209.7	5206	208.979	152.779	0.0012	0.0011	0.9568
MER31-int LTR ERV1	250.8	149	1.692	0	0.0010	0.0010	0.9567
Charlie4a DNA hAT-Charlie	209.2	1121	24.536	13.233	0.0011	0.0011	0.9564
MER58D DNA hAT-Charlie	126.9	292	1.692	0	0.0010	0.0010	0.9563
RLTR13A LTR ERVK	658.4	211	7.615	1.203	0.0011	0.0010	0.9562
L1M4c LINE L1	303.7	2049	104.067	72.179	0.0012	0.0011	0.9561
LTR69 LTR ERV1	248.5	74	0.846	0	0.0010	0.0010	0.9560
B2_Mm2 SINE B2	153.7	90007	5896.257	5027.258	0.0014	0.0014	0.9560
L1MCc LINE L1	226.7	529	11.845	6.015	0.0011	0.0011	0.9558
MamRep137 DNA TcMar-Tigger	162.4	339	5.076	2.406	0.0011	0.0010	0.9556
MLTR18B_MM LTR ERVK	359.5	468	112.527	99.847	0.0017	0.0016	0.9548
LTR33A_ LTR ERV1	201.4	504	38.919	32.48	0.0014	0.0013	0.9541
MER47B DNA TcMar-Tigger	210.2	509	15.229	9.624	0.0011	0.0011	0.9541
MT2C_Mm LTR ERV1	362.5	1982	88.837	51.728	0.0011	0.0011	0.9540
RLTR9F LTR ERVK	342.7	411	11.845	4.812	0.0011	0.0010	0.9539
UON26 Unknown Unknown	161.9	108	0.846	0	0.0010	0.0010	0.9538
RLTR20C2_MM LTR ERVK	125.8	3620	109.143	83.006	0.0012	0.0012	0.9537
LTR107_Mam LTR LTR	185.9	93	0.846	0	0.0010	0.0010	0.9533
MT2B LTR ERV1	127.4	13977	906.14	780.735	0.0015	0.0014	0.9533
Charlie13a DNA hAT-Charlie	236.4	184	5.922	3.609	0.0011	0.0011	0.9532
L1MC2 LINE L1	285.4	2252	84.607	50.525	0.0011	0.0011	0.9531
BGLII_C LTR ERVK	231.5	74	0.846	0	0.0010	0.0010	0.9529
MLT2B3 LTR ERV1	242.1	1483	27.92	9.624	0.0011	0.0010	0.9527
Kangal1a DNA TcMar-Tc2	227.7	187	3.384	1.203	0.0011	0.0010	0.9525
RLTR18B LTR ERVK	320.8	989	60.071	42.104	0.0012	0.0011	0.9524
RLTR13B4 LTR ERVK	739.7	412	20.306	4.812	0.0011	0.0010	0.9523
PB1D10 SINE Alu	95.1	65054	4439.326	3932.545	0.0017	0.0016	0.9523
LTRIS4 LTR ERV1	352	147	17.767	14.436	0.0013	0.0013	0.9521
ERV3-16A3_I-int LTR ERV1	261	908	39.765	26.466	0.0012	0.0011	0.9519
LTR81C LTR Gypsy	181	99	16.075	14.436	0.0019	0.0018	0.9518
B1_Mur2 SINE Alu	122.1	31869	2621.123	2306.115	0.0017	0.0016	0.9516
RLTR16 LTR ERVK	317.2	1350	111.681	85.412	0.0013	0.0012	0.9513
RLTR9A3A LTR ERVK	296.8	222	3.384	0	0.0011	0.0010	0.9512



## Appendix

Tigger17a DNA TcMar-Tigger	310	908	42.303	26.466	0.0012	0.0011	0.9511
ID4 SINE ID	70.8	23970	1264.874	1119.975	0.0017	0.0017	0.9511
MER94 DNA hAT-Blackjack	105.5	466	2.538	0	0.0011	0.0010	0.9509
MamGypLTR2c LTR Gypsy	148.8	110	0.846	0	0.0011	0.0010	0.9509
L1MB4 LINE L1	251	2690	143.832	103.456	0.0012	0.0012	0.9507
Mam_R4 LINE Dong-R4	176.5	138	2.538	1.203	0.0011	0.0010	0.9504
RLTR30D2_MM LTR ERV1	352.6	741	61.763	45.713	0.0012	0.0012	0.9503
ID4_ SINE ID	70	26384	1098.198	951.558	0.0016	0.0015	0.9502
RLTR16B_MM LTR ERVK	375.8	635	25.382	12.03	0.0011	0.0011	0.9494
Cheshire DNA hAT-Charlie	195.2	1091	22.844	10.827	0.0011	0.0011	0.9490
MER90 LTR ERV1	223.7	245	4.23	1.203	0.0011	0.0010	0.9487
ORR1E LTR ERV L-MaLR	235.1	22742	884.989	564.198	0.0012	0.0011	0.9485
RLTR13D3 LTR ERVK	640.5	339	24.536	12.03	0.0011	0.0011	0.9483
Charlie7 DNA hAT-Charlie	248	844	43.15	30.075	0.0012	0.0011	0.9482
MLT1E1 LTR ERV L-MaLR	214	576	21.998	14.436	0.0012	0.0011	0.9479
RLTR48C LTR ERV1	405.1	149	8.461	4.812	0.0011	0.0011	0.9470
RLTR18 LTR ERVK	395.2	610	36.381	21.654	0.0012	0.0011	0.9469
RLTR20A4 LTR ERVK	78.7	11527	306.277	241.799	0.0013	0.0013	0.9469
X3_LINE LINE RTE-BovB	85.7	260	10.153	8.421	0.0015	0.0014	0.9466
L1ME4c LINE L1	190.3	968	115.912	99.847	0.0016	0.0015	0.9465
Tigger12A DNA TcMar-Tigger	173.6	86	0.846	0	0.0011	0.0010	0.9464
ID2 SINE ID	70.7	5671	293.586	256.235	0.0017	0.0016	0.9462
URR1B DNA hAT-Charlie	173.8	16889	956.058	745.848	0.0013	0.0013	0.9460
RLTR20A2 LTR ERVK	280	368	21.152	14.436	0.0012	0.0011	0.9459
Kanga1 DNA TcMar-Tc2	206.7	321	5.076	1.203	0.0011	0.0010	0.9458
Charlie9 DNA hAT-Charlie	127	232	1.692	0	0.0011	0.0010	0.9457
RLTR20A3_MM LTR ERVK	290.7	880	43.996	27.669	0.0012	0.0011	0.9455
L1ME3G LINE L1	177.6	1161	130.295	111.877	0.0016	0.0015	0.9453
L1MC5 LINE L1	227	3476	176.828	123.907	0.0012	0.0012	0.9452
LTRIS_Mus LTR ERV1	508.2	919	170.906	135.937	0.0014	0.0013	0.9452
LTR84b LTR ERV L	192.1	187	3.384	1.203	0.0011	0.0010	0.9445
MTE2a-int LTR ERV L-MaLR	441.9	795	32.151	10.827	0.0011	0.0010	0.9444
L1ME3E LINE L1	238.3	327	21.152	15.639	0.0013	0.0012	0.9444
MTEb-int LTR ERV L-MaLR	458.6	635	46.534	27.669	0.0012	0.0011	0.9441
B1F SINE Alu	109.8	45170	2567.82	2146.118	0.0015	0.0014	0.9440
RLTR18-int LTR ERVK	322.2	506	16.075	6.015	0.0011	0.0010	0.9438
L2c LINE L2	126.5	12298	675.163	549.762	0.0014	0.0014	0.9438
RLTR9A3 LTR ERVK	283.2	774	24.536	10.827	0.0011	0.0010	0.9438
LTR48 LTR ERV1	163.1	651	10.153	3.609	0.0011	0.0010	0.9437
MER97b DNA hAT-Tip100	247.2	55	8.461	7.218	0.0016	0.0015	0.9436
LTR33B LTR ERV L	180.7	268	16.921	13.233	0.0013	0.0013	0.9436
Charlie1a DNA hAT-Charlie	278	2660	181.059	128.719	0.0012	0.0012	0.9431
LTRIS6 LTR ERV1	281.6	149	2.538	0	0.0011	0.0010	0.9430
L1ME3F LINE L1	247.9	308	5.922	1.203	0.0011	0.0010	0.9426
MER53 DNA hAT	121.6	1693	32.997	19.248	0.0012	0.0011	0.9424
RLTR13D3A1 LTR ERVK	478.7	613	62.609	42.104	0.0012	0.0011	0.9424
RLTR9B LTR ERVK	333	206	9.307	4.812	0.0011	0.0011	0.9423
RMER17B2 LTR ERVK	474.3	898	33.843	7.218	0.0011	0.0010	0.9421
L1M1 LINE L1	146.5	881	11.845	3.609	0.0011	0.0010	0.9416
MLT1A0 LTR ERV L-MaLR	201.8	10413	497.489	344.053	0.0012	0.0012	0.9410
Charlie15b DNA hAT-Charlie	198	272	3.384	0	0.0011	0.0010	0.9409
LTR65 LTR ERV1	202.8	162	8.461	6.015	0.0013	0.0012	0.9408
LTR31 LTR ERV1	214.2	187	2.538	0	0.0011	0.0010	0.9404
MTD-int LTR ERV L-MaLR	454.5	2521	227.593	144.358	0.0012	0.0011	0.9394
ID SINE ID	59.3	8022	293.586	246.611	0.0016	0.0015	0.9389
MER67A LTR ERV1	221.6	176	2.538	0	0.0011	0.0010	0.9389
ORR1G-int LTR ERV L-MaLR	371.2	70	1.692	0	0.0011	0.0010	0.9389
MLT1E LTR ERV L-MaLR	216.8	534	15.229	7.218	0.0011	0.0011	0.9388
MYSERV-int LTR ERVK	448.9	1435	117.604	70.976	0.0012	0.0011	0.9388
MamRep38 DNA hAT	140.9	212	18.614	15.639	0.0016	0.0015	0.9386

## Appendix

MER91A DNA hAT-Tip100	115.6	370	19.46	15.639	0.0015	0.0014	0.9386
MER5A1 DNA hAT-Charlie	113.7	3905	62.609	31.278	0.0011	0.0011	0.9382
Tigger6a DNA TcMar-Tigger	309.9	369	10.153	2.406	0.0011	0.0010	0.9378
RLTR20A LTR ERVK	233.6	361	28.766	21.654	0.0013	0.0013	0.9371
LTR84a LTR ERV L	185.2	136	1.692	0	0.0011	0.0010	0.9371
ERV B7_2-LTR_MM LTR ERVK	302.3	425	27.92	18.045	0.0012	0.0011	0.9369
MER44B DNA TcMar-Tigger	225.3	1493	62.609	37.292	0.0012	0.0011	0.9365
MER20 DNA hAT-Charlie	147.1	5994	365.502	286.31	0.0014	0.0013	0.9365
RLTR11D LTR ERVK	347.5	529	25.382	12.03	0.0011	0.0011	0.9362
MLTR31F_MM LTR ERVK	362.7	403	20.306	9.624	0.0011	0.0011	0.9358
LTR81A LTR Gypsy	168	180	3.384	1.203	0.0011	0.0010	0.9351
MLT1J LTR ERV L-MaLR	182.8	2943	142.986	98.644	0.0013	0.0012	0.9349
UCON4 Unknown Unknown	142.6	72	40.611	37.292	0.0050	0.0046	0.9348
BGLII_B LTR ERVK	354.4	960	77.838	50.525	0.0012	0.0011	0.9347
RMER12C LTR ERVK	187	4954	558.406	460.742	0.0016	0.0015	0.9342
MIR3 SINE MIR	98.4	16085	851.992	689.308	0.0015	0.0014	0.9332
RLTR12C LTR ERVK	295.5	319	85.453	73.382	0.0019	0.0018	0.9328
Tigger6b DNA TcMar-Tigger	216.2	217	3.384	0	0.0011	0.0010	0.9327
MER2B DNA TcMar-Tigger	122.1	1093	35.535	24.06	0.0013	0.0012	0.9321
ORR1G LTR ERV L-MaLR	193.4	4680	175.982	102.253	0.0012	0.0011	0.9318
MLTR12 LTR ERV1	247.8	477	17.767	8.421	0.0012	0.0011	0.9313
IAPEY5_LTR LTR ERVK	285.5	237	7.615	2.406	0.0011	0.0010	0.9308
MER3 DNA hAT-Charlie	133.1	2967	188.673	147.967	0.0015	0.0014	0.9302
RLTR9A LTR ERVK	308.1	648	25.382	9.624	0.0011	0.0010	0.9300
IAPLTR2a2_Mm LTR ERVK	407	1541	124.372	70.976	0.0012	0.0011	0.9290
MLTR13 LTR ERVK	307.1	205	17.767	12.03	0.0013	0.0012	0.9289
FordPrefect_a DNA hAT-Tip100	184.1	60	0.846	0	0.0011	0.0010	0.9289
MER45A DNA hAT-Tip100	121.6	769	8.461	1.203	0.0011	0.0010	0.9288
ERV B4_1B-LTR_MM LTR ERVK	472.9	648	75.3	48.119	0.0012	0.0012	0.9288
RLTR49 LTR ERVK	295.2	933	38.073	15.639	0.0011	0.0011	0.9284
LTR101_Mam LTR ERV L	240.1	157	4.23	1.203	0.0011	0.0010	0.9278
MurSatRep1 Unknown Unknown	353.6	2600	467.876	366.909	0.0015	0.0014	0.9272
L3b LINE CR1	109.5	2304	113.373	86.615	0.0014	0.0013	0.9268
L1MEi LINE L1	215.1	478	21.152	12.03	0.0012	0.0011	0.9264
MER5C1 DNA hAT-Charlie	140.6	75	0.846	0	0.0011	0.0010	0.9257
Tigger2b DNA TcMar-Tigger	237.2	176	3.384	0	0.0011	0.0010	0.9250
PB1D7 SINE Alu	96.5	24346	2132.095	1792.442	0.0019	0.0018	0.9242
B1F2 SINE Alu	110.9	16513	1382.477	1134.411	0.0018	0.0016	0.9228
MLT1O LTR ERV L-MaLR	172	538	16.921	8.421	0.0012	0.0011	0.9223
RLTR12D LTR ERVK	313.4	647	40.611	21.654	0.0012	0.0011	0.9221
BC1_Mm scRNA scRNA	81	6758	860.453	750.66	0.0026	0.0024	0.9220
RLTR10A LTR ERVK	348.3	800	86.299	57.743	0.0013	0.0012	0.9218
RLTR41C LTR ERV1	291.4	157	31.305	25.263	0.0017	0.0016	0.9216
hAT-N1_Mam DNA hAT?	160.1	185	2.538	0	0.0011	0.0010	0.9211
L4_B_Mam LINE RTE-X	163.4	361	5.076	0	0.0011	0.0010	0.9208
SRV_MM-int LTR ERVK	757.2	129	20.306	10.827	0.0012	0.0011	0.9197
MLT1E1A-int LTR ERV L-MaLR	344.3	28	0.846	0	0.0011	0.0010	0.9193
MER96 DNA hAT-Tip100	117.4	104	7.615	6.015	0.0016	0.0015	0.9193
MER31A LTR ERV1	239.2	707	34.689	18.045	0.0012	0.0011	0.9183
Tigger15a DNA TcMar-Tigger	180.4	704	55.841	40.901	0.0014	0.0013	0.9183
ERV B5_1-LTR_MM LTR ERVK	377.2	143	12.691	7.218	0.0012	0.0011	0.9179
SSU-rRNA_Hsa rRNA rRNA	322.1	45	3584.796	3284.139	0.2483	0.2276	0.9165
LTR28B LTR ERV1	117	144	30.458	26.466	0.0028	0.0026	0.9156
UCON6 Unknown Unknown	147.8	45	100.682	91.427	0.0161	0.0147	0.9138
MLTR32C_MM LTR ERVK	282.3	127	3.384	0	0.0011	0.0010	0.9138
Eulor9A DNA DNA	147.4	60	0.846	0	0.0011	0.0010	0.9127
LTR81 LTR Gypsy	163.3	70	5.076	3.609	0.0014	0.0013	0.9111
RLTR12B2 LTR ERVK	363	95	3.384	0	0.0011	0.0010	0.9106
Charlie23a DNA hAT-Charlie	157.9	157	5.076	2.406	0.0012	0.0011	0.9106
Kanga1d DNA TcMar-Tc2	160.1	131	3.384	1.203	0.0012	0.0011	0.9105

## Appendix

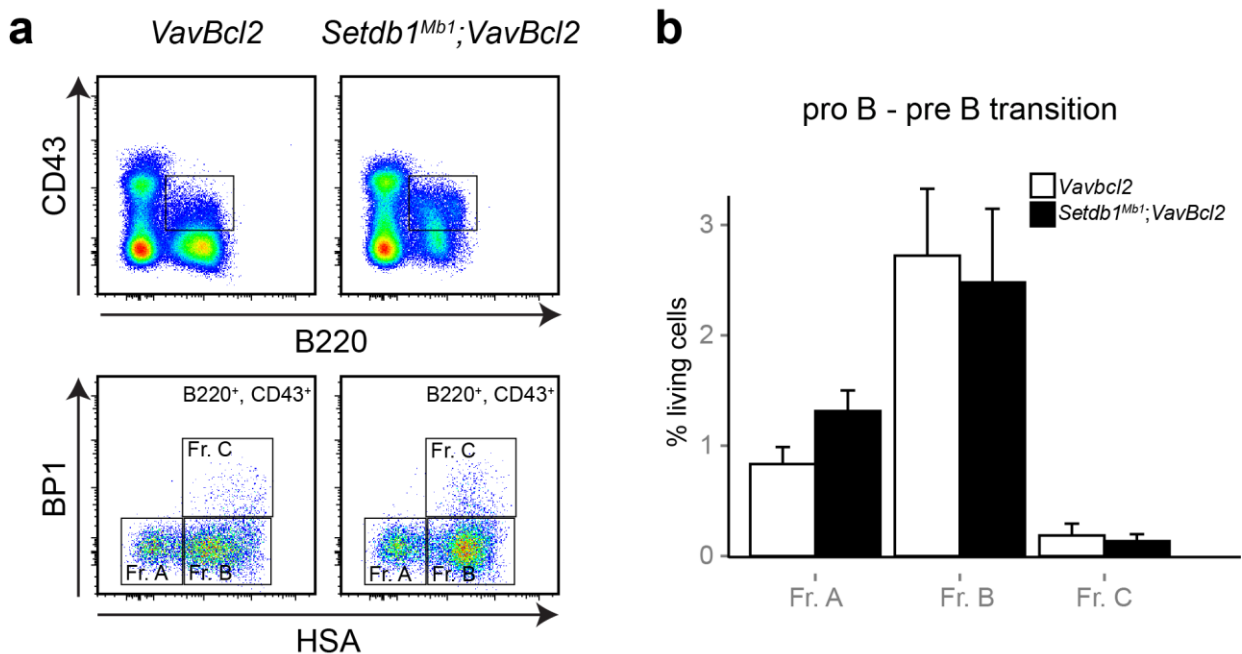
MER65D LTR ERV1	169.3	101	1.692	0	0.0011	0.0010	0.9100
IAPA_MM-int LTR ERVK	296.3	126	16.921	12.03	0.0015	0.0013	0.9099
LTR89 LTR ERV1?	175.1	146	2.538	0	0.0011	0.0010	0.9097
LTR48B LTR ERV1	209.1	302	7.615	1.203	0.0011	0.0010	0.9094
Charlie10 DNA hAT-Charlie	206.7	371	16.921	8.421	0.0012	0.0011	0.9092
LTRIS_Mm LTR ERV1	452.9	555	225.054	181.65	0.0019	0.0017	0.9089
MLT1N2 LTR ERV1-MaLR	162.8	692	19.46	7.218	0.0012	0.0011	0.9073
MTB_Mm-int LTR ERV1-MaLR	647.1	379	71.916	42.104	0.0013	0.0012	0.9060
MLT1G1-int LTR ERV1-MaLR	319.7	51	1.692	0	0.0011	0.0010	0.9060
UCON13 LINE? Penelope?	193.5	42	0.846	0	0.0011	0.0010	0.9057
Charlie15a DNA hAT-Charlie	123.7	217	25.382	20.451	0.0019	0.0018	0.9056
RLTR31B_Mm LTR ERVK	361.4	349	39.765	24.06	0.0013	0.0012	0.9053
MamRep605 LTR? LTR?	185.5	629	13.537	1.203	0.0011	0.0010	0.9053
L1MEb LINE L1	219.7	352	25.382	15.639	0.0013	0.0012	0.9051
LTR80B LTR ERV1	208.3	116	2.538	0	0.0011	0.0010	0.9049
MER117 DNA hAT-Charlie	122	745	60.071	45.713	0.0017	0.0015	0.9049
MER102c DNA hAT-Charlie	161.5	363	19.46	12.03	0.0013	0.0012	0.9048
LTR86A1 LTR ERV1	174.7	92	1.692	0	0.0011	0.0010	0.9048
MLTR31C_MM LTR ERVK	389	227	18.614	8.421	0.0012	0.0011	0.9047
MER97a DNA hAT-Tip100	302.9	53	1.692	0	0.0011	0.0010	0.9047
MER81 DNA hAT-Blackjack	89.2	512	10.153	4.812	0.0012	0.0011	0.9043
MamGypLTR1c LTR Gypsy	164.8	145	2.538	0	0.0011	0.0010	0.9040
MT2B2 LTR ERV1	188	3655	522.025	405.405	0.0018	0.0016	0.9036
RLTR20C1_MM LTR ERVK	82	5561	566.02	466.757	0.0022	0.0020	0.9029
MER135 DNA DNA	109.3	287	3.384	0	0.0011	0.0010	0.9026
RLTR40 LTR ERVK	339.1	2651	231.823	120.298	0.0013	0.0011	0.9014
L1M2a LINE L1	277.7	208	14.383	7.218	0.0012	0.0011	0.9007
UCON14 DNA? DNA?	115.7	66	0.846	0	0.0011	0.0010	0.9003
Ricksha DNA MuDR	265.2	85	2.538	0	0.0011	0.0010	0.8988
MER92B LTR ERV1	241.8	238	11.845	4.812	0.0012	0.0011	0.8987
RMER16B LTR ERVK	355.4	84	3.384	0	0.0011	0.0010	0.8982
RMER17C LTR ERVK	260.9	6479	494.951	271.874	0.0013	0.0012	0.8979
MLT1G LTR ERV1-MaLR	219.2	896	227.593	184.056	0.0022	0.0019	0.8973
MLT1G-int LTR ERV1-MaLR	266.2	39	2.538	1.203	0.0012	0.0011	0.8967
5S rRNA rRNA	80.5	1038	24.536	13.233	0.0013	0.0012	0.8954
7SK RNA RNA	170.6	691	38.073	21.654	0.0013	0.0012	0.8947
HERVL40-int LTR ERV1	227.9	204	10.999	4.812	0.0012	0.0011	0.8924
LTR83 LTR ERV1	196.7	161	9.307	4.812	0.0013	0.0012	0.8903
MLT-int LTR ERV1-MaLR	281.1	846	65.993	32.48	0.0013	0.0011	0.8897
X7C_LINE LINE CR1	95.9	155	5.922	3.609	0.0014	0.0012	0.8887
Charlie26a DNA hAT-Charlie	164.9	98	10.153	7.218	0.0016	0.0014	0.8885
RLTR31D_MM LTR ERVK	121.9	8270	488.182	321.196	0.0015	0.0013	0.8884
MER54B LTR ERV1	270.2	114	9.307	4.812	0.0013	0.0012	0.8879
Charlie1b DNA hAT-Charlie	194.7	952	46.534	20.451	0.0013	0.0011	0.8875
Charlie4 DNA hAT-Charlie	165	161	3.384	0	0.0011	0.0010	0.8870
Ricksha_0 DNA MuDR	242.6	82	2.538	0	0.0011	0.0010	0.8869
RLTR41B LTR ERV1	295.8	98	5.076	1.203	0.0012	0.0010	0.8863
L1MEg2 LINE L1	199.8	112	4.23	1.203	0.0012	0.0011	0.8862
PB1D11 SINE Alu	83.6	18484	1176.036	863.74	0.0018	0.0016	0.8852
RLTR30B_MM LTR ERV1	352.7	283	21.152	7.218	0.0012	0.0011	0.8848
MER105 DNA DNA	128.1	171	4.23	1.203	0.0012	0.0011	0.8842
MER123 DNA DNA	156.9	41	0.846	0	0.0011	0.0010	0.8838
FordPrefect DNA hAT-Tip100	290.1	291	95.606	74.585	0.0021	0.0019	0.8832
LTR50 LTR ERV1	183	549	69.378	49.322	0.0017	0.0015	0.8819
UCON28c Unknown Unknown	168.9	37	0.846	0	0.0011	0.0010	0.8808
ERV5_2-LTR_MM LTR ERVK	397.3	520	92.222	56.54	0.0014	0.0013	0.8806
LTR58 LTR ERV1	105	9	4.23	3.609	0.0055	0.0048	0.8800
Charlie10b DNA hAT-Charlie	128.8	48	0.846	0	0.0011	0.0010	0.8796
MER106A DNA hAT-Charlie	160.5	115	2.538	0	0.0011	0.0010	0.8791
MERX DNA TcMar-Tigger	190.9	64	1.692	0	0.0011	0.0010	0.8784

## Appendix

MLTR31E_MM LTR ERVK	362.9	164	10.999	2.406	0.0012	0.0010	0.8781
MamTip1 DNA hAT-Tip100	147.9	82	30.458	25.263	0.0035	0.0031	0.8780
Arthur1C DNA hAT-Tip100	151.9	155	10.153	6.015	0.0014	0.0013	0.8772
MER47A DNA TcMar-Tigger	189.9	127	3.384	0	0.0011	0.0010	0.8770
RLTR12F LTR ERVK	296.4	60	2.538	0	0.0011	0.0010	0.8751
HAL1b LINE L1	194	763	40.611	16.842	0.0013	0.0011	0.8740
RLTR34D_MM LTR ERVK	375.4	410	60.917	33.683	0.0014	0.0012	0.8732
RLTR34C_MM LTR ERVK	365.3	402	50.764	25.263	0.0013	0.0012	0.8710
Charlie13b DNA hAT-Charlie	204.7	114	7.615	3.609	0.0013	0.0012	0.8706
4.5SRNA scRNA scRNA	69.9	1517	611.708	518.485	0.0068	0.0059	0.8701
HAL1M8 LINE L1	211.5	710	55.841	28.872	0.0014	0.0012	0.8691
LTR109A2 LTR ERV1	150.7	37	0.846	0	0.0012	0.0010	0.8683
Charlie1 DNA hAT-Charlie	283.1	977	109.143	57.743	0.0014	0.0012	0.8667
RLTR53B_Mm LTR ERVK	195	160	5.076	0	0.0012	0.0010	0.8601
RLTR14 LTR ERV1	258.1	2341	221.67	105.862	0.0014	0.0012	0.8598
RMER16_Mm LTR ERVK	354.1	435	44.842	16.842	0.0013	0.0011	0.8592
MLTR25A LTR ERVK	424	2168	417.112	228.566	0.0015	0.0012	0.8589
MLT2E LTR ERVL	189	126	6.769	2.406	0.0013	0.0011	0.8573
L1MEh LINE L1	198.1	202	6.769	0	0.0012	0.0010	0.8553
MER131 SINE? SINE?	109.3	274	5.076	0	0.0012	0.0010	0.8551
L1M LINE L1	101.3	459	9.307	1.203	0.0012	0.0010	0.8548
RLTR26B_MM LTR ERVK	358.9	507	35.535	3.609	0.0012	0.0010	0.8532
ETnERV2-int LTR ERVK	840.8	1805	603.247	291.121	0.0014	0.0012	0.8528
MLT1F2-int LTR ERVL-MaLR	322.8	68	15.229	9.624	0.0017	0.0014	0.8492
LTR108b_Mam LTR ERVL	250.5	19	0.846	0	0.0012	0.0010	0.8491
MLT1B-int LTR ERVL-MaLR	255.3	165	11.845	3.609	0.0013	0.0011	0.8474
MamTip3 DNA hAT-Tip100	155.5	180	9.307	3.609	0.0013	0.0011	0.8472
UCON31 Unknown Unknown	156.4	89	2.538	0	0.0012	0.0010	0.8458
MER5C DNA hAT-Charlie	164.5	242	10.153	2.406	0.0013	0.0011	0.8449
X7A_LINE LINE CR1	100	159	38.919	30.075	0.0034	0.0029	0.8387
Charlie8 DNA hAT-Charlie	175.3	728	191.212	139.546	0.0025	0.0021	0.8380
LTR81B LTR Gypsy	191.3	204	23.69	13.233	0.0016	0.0013	0.8333
Charlie14a DNA hAT-Charlie	172.7	72	2.538	0	0.0012	0.0010	0.8305
YREP_Mm Unknown Unknown	413.1	2800	644.705	327.211	0.0016	0.0013	0.8238
MER130 Unknown Unknown	177.1	89	3.384	0	0.0012	0.0010	0.8233
X8_LINE LINE CR1	119.7	158	8.461	3.609	0.0014	0.0012	0.8227
Charlie17b DNA hAT-Charlie	252.9	130	42.303	28.872	0.0023	0.0019	0.8213
MER113A DNA hAT-Charlie	155.1	140	6.769	1.203	0.0013	0.0011	0.8046
DNA1_Mam DNA TcMar	153.4	45	1.692	0	0.0012	0.0010	0.8031
MER57D LTR ERV1	209.8	300	246.206	185.259	0.0049	0.0039	0.8029
Tigger10 DNA TcMar-Tigger	214.9	242	48.226	27.669	0.0019	0.0015	0.7949
L1MEj LINE L1	225.8	178	19.46	7.218	0.0015	0.0012	0.7948
X6B_LINE LINE CR1	93.9	274	29.612	18.045	0.0022	0.0017	0.7910
LTR104_Mam LTR Gypsy	174.6	254	63.455	40.901	0.0024	0.0019	0.7908
LTR40c LTR ERVL	211	248	17.767	2.406	0.0013	0.0010	0.7809
LTRIS3 LTR ERV1	379.5	165	25.382	6.015	0.0014	0.0011	0.7799
Charlie19a DNA hAT-Charlie	181.4	240	16.921	3.609	0.0014	0.0011	0.7798
LTR108d_Mam LTR ERVL	198.3	30	1.692	0	0.0013	0.0010	0.7786
MLTR25C LTR ERVK	559.8	504	137.909	44.51	0.0015	0.0012	0.7776
X5B_LINE LINE CR1	73.1	79	1.692	0	0.0013	0.0010	0.7734
HAL1_SS LINE L1	259.7	79	9.307	2.406	0.0015	0.0011	0.7686
LTR45 LTR ERV1	103.3	31	2.538	1.203	0.0018	0.0014	0.7674
LTR53B LTR ERVL	247.6	83	10.999	3.609	0.0015	0.0012	0.7658
MER97d DNA hAT-Tip100	193.2	28	1.692	0	0.0013	0.0010	0.7617
Eulor2B DNA? DNA?	145	42	5.076	2.406	0.0018	0.0014	0.7609
RLTR12A LTR ERVK	352.1	829	228.439	102.253	0.0018	0.0014	0.7575
Tigger1a_Art DNA TcMar-Tigger	137.1	37	1.692	0	0.0013	0.0010	0.7499
LTR16D1 LTR ERVL	192.5	90	21.998	12.03	0.0023	0.0017	0.7465
Charlie4z DNA hAT-Charlie	114.2	431	21.998	3.609	0.0014	0.0011	0.7418
MER87 LTR ERV1	104.1	46	1.692	0	0.0014	0.0010	0.7389

## Appendix

Tigger12c DNA TcMar-Tigger	169.8	97	7.615	1.203	0.0015	0.0011	0.7338
MER129 LTR? LTR?	182.4	63	4.23	0	0.0014	0.0010	0.7309
FAM SINE Alu	60	117	7.615	3.609	0.0021	0.0015	0.7263
Zaphod2 DNA hAT-Tip100	238.4	169	31.305	9.624	0.0018	0.0012	0.6972
Eulor5A DNA? DNA?	154.2	80	7.615	1.203	0.0016	0.0011	0.6786
LTR88a LTR Gypsy?	146.8	64	11.845	4.812	0.0023	0.0015	0.6689
MER66C LTR ERV1	85.5	19	0.846	0	0.0015	0.0010	0.6576
LTR108a_Mam LTR ERV1	203.2	39	11.845	4.812	0.0025	0.0016	0.6443
UCON23 DNA hAT?	121.7	22	1.692	0	0.0016	0.0010	0.6128
MER50-int LTR ERV1	184.5	14	1.692	0	0.0017	0.0010	0.6042
Arthur1A DNA hAT-Tip100	127.7	148	17.767	1.203	0.0019	0.0011	0.5483
X6A_LINE LINE CR1	103.3	145	16.921	2.406	0.0021	0.0012	0.5450
MER34B-int LTR ERV1	204.9	81	40.611	14.436	0.0034	0.0019	0.5425
HY3 scRNA scRNA	65.5	15	0.846	0	0.0019	0.0010	0.5373
LTR9 LTR ERV1	112.2	13	1.692	0	0.0022	0.0010	0.4630
LTR28 LTR ERV1	107.6	10	4.23	1.203	0.0049	0.0021	0.4295
Eulor6E DNA? DNA?	134.2	18	3.384	0	0.0024	0.0010	0.4165
Ricksha_b DNA MuDR	182.5	2	1.692	0	0.0056	0.0010	0.1774
LTR44 LTR ERV1	71.3	3	1.692	0	0.0089	0.0010	0.1122
MER50B LTR ERV1	91.4	14	11.845	0	0.0103	0.0010	0.0975



**Figure 5.1 Enforced *Bcl2* expression partially rescues B cell development in *Setdb1<sup>Mb1</sup>* pro B cells**  
**a)** *Setdb1<sup>Mb1</sup>; VavBcl2* and *VavBcl2* bone marrow B cells stained with the Hardy scheme. **b)** Bargraphs depicting average and standard deviation of 6 biological replicates.

## 6 Material and methods

### 6.1 Materials

#### 6.1.1 Mice

Transgenic mouse	Strain
<i>Wild type (CD45.2)</i>	C57BL/6
<i>Suv4-20h2ko</i>	C57BL/6
<i>Setdb1 flox; flox</i>	C57BL/6
<i>Mb1cre</i>	C57BL/6
<i>VavBcl2</i> mice	C57BL/6
<i>Vavcre</i>	C57BL/6
<i>Wild type (CD45.1)</i>	B6/SJL

#### 6.1.2 Cell lines

Mouse embryonic fibroblasts (MEFs)

OP-9 (stromal cells)

#### 6.1.3 Antibodies and dyes

ANTIBODY	CLONE	FLUOROCHROME	DILUTION	PROVIDER
CD45R/B220	RA3-6B2	PE	1/1000	Pharmigen
CD138/Synd	281-2	Biotinylated	1/1000	Pharmigen
IgD	11-26c.2a	FITC	1/250	Pharmigen
CD24 (HSA)	M1/69	FITC	1/500	Pharmigen
CD11b/Mac-1	M1/70	PE		Pharmigen
CD25	PC61.5	APC-AlexaFluor750	1/500	eBioscience
Ter119	TER119	APC-AlexaFluor750	1/500	eBioscience
CD8a	53-6.7	PE-Cy7	1/2501/1000	eBioscience
CD45.1	A20	PE-Cy7	1/250	eBioscience
CD5	53-7.3	PE	1/500	Pharmigen
CD19	1D3	PE	1/1000	Pharmigen
CD11c	HL3	PE	1/250	Pharmigen
Sca-1	D7	FITC	1/500	Pharmigen
CD4	(L3T4)(PM4-5)	PE	1/5000	Pharmigen
CD8a	53-6.7	PE	1/1000	Pharmigen
CD95R	Jo2	PE	1/1000	Pharmigen
CD25	PC61	PE-Cy5	1/500	eBioscience
CD45.2	104	APC	1/250	eBioscience
IgM	II/41	FITC	1/250	Pharmigen
CD117	2B8	APC-Alexa Fluor780	1/500	Pharmigen
IgM	II/41	PE-Cy7	1/250	eBioscience
CD34	RAM34	Alexa-Fluor 647	1/200	eBioscience
CD45R/B220	RA3-SB2	APC-Alexa-Fluor750	1/250	eBioscience
CD21/CD35	7G6	APC	1/500	Pharmigen
CD8a	53-6.7	PE-Cy7	1/1000	eBioscience
Ly-6G(Gr-1)	RB6-8C5	PE	1/5000	eBioscience
CD19	1D3	APC	1/500	Pharmigen
CD19	1D3	APC-Cy7	1/1000	Pharmigen
CD117	2B8	PE	1/500	Pharmigen
Streptavidin		APC	1/1000	Pharmigen
CD3e	145-2C11	PE	1/250	Pharmigen
Streptavidin		APC-Cy7	1/1000	Pharmigen
Streptavidin		PE-Cy7	1/1000	Pharmigen
Streptavidin		PE-Cy5	1/1000	Pharmigen
Streptavidin		PE	1/1000	Pharmigen
PNA		FITC	1/500	Vector



## Materials and methods

<b>Cd11b/Mac1-<math>\alpha</math></b>	M1/70	APC-Cy7	1/4000	eBioscience
<b>CD44</b>	IM7	FITC	1/1000	eBioscience
<b>CD115</b>	AFS98	APC	1/3500	eBioscience
<b>Ly6-C</b>	HK1.4	PE-Cy5	1/2500	eBioscience
<b>DX-5</b>	DX5	PE	1/1000	eBioscience
<b>Ter119</b>	Ter-119	PE	1/1000	eBioscience
<b>CD127(IL7R<math>\alpha</math>)</b>	A7R34	PE-Cy5	1/500	eBioscience
<b>F4/80</b>	BM8	PE	1/1000	Biolegend
<b>CD62L</b>	MEL-14	FITC	1/1000	eBioscience
<b>CD16/32</b>	93	PE-Cy7	1/1000	eBioscience
<b>CD25</b>	BC96	APC	1/500	Biolegend
<b>BP-1</b>	6C3	PE	1/200	eBioscience
<b>CD43</b>	S7	APC	1/500	Pharmingen
<b>CD40</b>	3/23	PE	1/500	Biolegend
<b>CD279 (PD-1)</b>	RMP1-30	Bio	1/1000	eBioscience
<b>CD278 (Icos)</b>	C398.4A	FITC	1/250	eBioscience
<b>CD278 (Icos)</b>	C398.4A	APC	1/250	eBioscience
<b>7-AAD</b>	none	PE-Cy5	1/100	eBioscience
<b>CD44</b>	IM7	PECy7	1/100	eBioscience
<b>CD4</b>	GK1.5	FITC	1/1000	eBioscience
<b>CD185(CXCR5)</b>	SPRCL5	PE	1/100	eBioscience
<b>CD34</b>	RAM34	eFluor660	1/100	eBioscience
<b>Fcblock(CD16/CD32)</b>	2.4G2	N/A	1/500	Pharmingen
<b><math>\alpha</math>-H3K9me3</b>		N/A		Active motif
<b><math>\alpha</math>-H3K4me3</b>		N/A		Diagenode
<b><math>\alpha</math>-H3K9ac</b>		N/A		
<b><math>\alpha</math>-Setdb1</b>		N/A		

### 6.1.4 Technical devices and material

Product	Provider
Axiovert 200 M	Zeiss
Optical microscope (DM IL)	Leica
Centrifuge (Haereus Multifuge 4KR)	Thermo Scientific
Banchtop centrifuge (Haereus Pico 17)	Thermo Scientific
Thermoblock/shaker	Eppendorf
Casy counter	Innovatis
Facs Canto	BD Bioscience
MoFlo	Beckman Coulter
Facs Aria	BD Bioscience
Polystyrene round-bottom tubes 12x75 mm (5ml)	Falcon
Tissue culture hoods (Hera cell)	Thermo Scientific
Scissors	F.T.S.
Rollermixer SRT6	STUART
Microtome	Thermo Scientific

## Materials and methods

Embedding blocks	Sigma Aldrich
Light cycler 480II	Roche
PCR machine	PeQLabBiotechnology/Applied Biosystems
Infrared lamp	Beurer
Tailveiner (small)	Biomedical Instruments
Cell strainer 40 and 70 $\mu$ m	Falcon
AutoMACS	Miltenyi Biotec
Quadro MACS	Miltenyi Biotec
Magnetic columns	Miltenyi Biotec
Syringes	B/Braun
Needles	B/Braun
Blunt end needles	Stemcell Technology
Syringes (colony forming assay)	Stemcell Technology
Syringes transplantation	BD Bioscience
Tissue culture dishes	Cell Star/Corning
Colony forming assay dishes	Stemcell Technology
Transilluminator	Syngene
Slides for microscopy	Roth
Cytospin	Thermo Scientific

### 6.1.5 Kits

Product	Provider
Lineage Cell Depletion Kit, mouse	Miltenyi Biotec
CD45R (B220) MicroBeads, mouse	Miltenyi Biotec
B Cell Isolation Kit, mouse	Miltenyi Biotec



## Materials and methods

Annexin V kit	eBioscience
Rneasy mini kit	QIAGEN
Rneasy Plus micro kit	QIAGEN
DNeasy Blood and tissue kit	QIAGEN

### 6.1.6 Tissue culture media, cytokines and immunostimulants

Product	Provider
RPMI	Gibco
DMEM	Sigma Aldrich/PAA
IMDM	Gibco
Methylcellulose 3434	Stemcell Technology
Methylcellulose 3630	Stemcell Technology
Serum	Sigma Aldrich/PAA
$\beta$ -mercaptoethanol	Sigma Aldrich
Non-essential aminoacids	Sigma Aldrich
Penicillin/Streptomycin	Sigma Aldrich
Interleukin-7 (IL-7)	Pepr Tech
LPS (lipopolysaccharide)	Sigma Aldrich
CpG	Axxora
Interleukin-4 (IL-4)	Pepr Tech
CD40	eBioscience

### 6.1.7 Reagents and buffers

Product	Provider
Formaldehyde	Roth
Formaldehyde ChIP	Sigma Aldrich
May-Grünwald	Roth

## Materials and methods

GIEMSA	Roth
BD PharmLyse (erythrocytes lysis)	BD Bioscience
PBS	Homemade
EDTA	Sigma Aldrich
Haematoxylin	Roth
Eosin	Roth
Ethidium bromide	Roth
TBA (Tris-borate-EDTA)	Homemade
FACS buffer	Homemade
MACS buffer	Homemade
Fast SYBR green master mix	Applied Biosystems
Superscript III kit	Invitrogen
RNasin® Plus	Promega
EGTA	Roth
Triton	Roth
SDS	Fluka
Na-deoxycholate	Sigma Aldrich
Protease inhibitor cocktail	Roche
Proteinase K	Bioline
Glycogen	Roche
RnaseA	Sigma Aldrich/ Roche
NaCl	Promega
Ethanol	Porlabo chemicals
Dynabeads	Applied technology

## 6.1.8 Oligonucleotides

Primer	Sequence	Experiment
MLV1-gag F	TCTTGCCACCGTAGTTACAG	qRT-PCR
MLV1-gag R	CCAGTGTCCTTTTCTTTGCAG	qRT-PCR
MLV1-env F	AGAGGCAGCCACAAAAACAG	qRT-PCR
MLV1-env R	ATTATAGGTGGCCCCCAGTTC	qRT-PCR
MLV5-gag F	AGCTCCAAAGAATCCGAAACG	qRT-PCR
MLV5-gag R	ATCTGTATCTGGCGGTTCCG	qRT-PCR
MLV5-pol F	ATCATAGGTGGCTCCCAGTTC	qRT-PCR
MLV5-pol R	AACCGAATGGCAGATCAAGC	qRT-PCR
MLV5-env F	GTCAAAGAGAACAGGGTCACC	qRT-PCR
MLV5-env R	CGGGTAAAAGGGCCAGCTG	qRT-PCR
MLV5-LTR F	ATCCTGTTTGGCCTCTGTCTC	qRT-PCR
MLV5-LTR R	ACGCCATTTTGCAAGGCATG	qRT-PCR
MLV8-gag F	TGACCCAGCGTCTCTTCTTG	qRT-PCR
MLV8-gag R	GGACCGCTTCTAAAAACATGGG	qRT-PCR
MLV8-LTR F	CCTAGTCCCGGTACTTTCCAG	qRT-PCR
MLV8-LTR R	ACTGCAGTAACGCCATCTTG	qRT-PCR
MLV8-env (spliced) F	CCAGGGACACCGACCCACCGT	qRT-PCR
MLV8-env (spliced) R	TAGTCGGTCCCGGTAGGCCTCG	qRT-PCR
AI506816 F	CCTGCTATGAAGGGGACAAAG	qRT-PCR
AI506816 R	ATCTTCGGAAGAGCAGTCAGTG	qRT-PCR
Fcgr2b F	GGAAGGACACTGCACCAGTC	qRT-PCR
Fcgr2b R	CCAGTGACAGCAGCCACAAT	qRT-PCR
Tubb3 F	GGCAACTATGTAGGGGACTCAG	qRT-PCR
Tubb3 R	ATGGTTCCAGGTTCCAAGTC	qRT-PCR
Gapdh F	TCAAGAAGGTGGTGAAGCAG	qRT-PCR
Gapdh R	GTTGAAGTCGCAGGAGACAA	qRT-PCR
Hprt F	ATGAGCGCAAGTTGAATCTG	qRT-PCR
Hprt R	CAGATGGCCACAGGACTAGA	qRT-PCR
MLV1-gag F	TCTTGCCACCGTAGTTACAG	ChIP-qPCR
MLV1-gag R	CCAGTGTCCTTTTCTTTGCAG	ChIP-qPCR
MLV5-gag F	AGCTCCAAAGAATCCGAAACG	ChIP-qPCR
MLV5-gag R	ATCTGTATCTGGCGGTTCCG	ChIP-qPCR
MLV8-gag F	TGACCCAGCGTCTCTTCTTG	ChIP-qPCR
MLV8-gag R	GGACCGCTTCTAAAAACATGGG	ChIP-qPCR
IAPs-gag F	AGCAGGTGAAGCCACTG	ChIP-qPCR
IAPs-gag R	CTTGCCACACTTAGAGC	ChIP-qPCR
IAPs-global F	CGGGTCGCGGTAATAAAGGT	ChIP-qPCR
IAPs-global R	ACTCTCGTTCCCCAGCTGAA	ChIP-qPCR
Tubb3-intron1 F	TTCTGACTCGCATTCCCATCC	ChIP-qPCR
Tubb3-intron2 R	GGCTTAAGTGGAACCTCAAAG	ChIP-qPCR
Def8-intron1 F	TGAGCCTTCGGTTTCACAAC	ChIP-qPCR
Def8-intron2 R	CAAAGCGCACCTCACATTTTC	ChIP-qPCR
H19 F	AGCTTTGAGTACCCAGGTTCA	ChIP-qPCR
H19 R	GCCTCTGCTTTTATGGCTATGG	ChIP-qPCR
Gapdh F	CCATCCCACGGCTCTGCAC	ChIP-qPCR
Gapdh R	GCAAGGCTTCCGTGCTCTCG	ChIP-qPCR

### 6.1.9 Software and databases

Flowjo

<http://www.immgen.org/>

## 6.2 Methods

### 6.2.1 Mice and cell lines

*Suv420h2 ko* mice were generated as previously described (Schotta et al., 2008). Mice carrying engineered *Setdb1* floxed (*Setdb1*<sup>f</sup>) or delta (*Setdb1*<sup>Δ</sup>) allele were purchased at EUCOMM. *Mblcre*, *VavBcl2*, *Vavcre* transgenic animals have been previously described (Georgiades et al., 2002; Hobeika et al., 2006; Ogilvy et al., 1999). Mouse colonies were housed in ventilated cages in the mouse facility at the Adolf Butenandt Institute, in agreement with EU regulations.

OP9 stromal cells were cultivated in IMDM medium, while splenic B cells and progenitors were expanded in RPMI medium (Gibco). Media were supplemented with 10% FSC, 1% non-essential amino acid, 1% penicillin streptomycin and 0.2% β-mercaptoethanol (Sigma). For progenitor cell short-term culture RPMI was supplemented with IL-7 (PeproTech). MethoCult 3630 used for Pre B cell differentiation was purchased from Stem Cell Technology.

### 6.2.2 Flow cytometry and cell sorting

Bone marrow cells were obtained flushing tibiae and femurs. These cells were then filtered through 70 μm cell strainer to eliminate debris and pieces of non-hematopoietic tissue. Hematopoietic cells were also taken from spleen and thymus gently smashing the organs on the surface of a 70 μm cell strained using a syringe piston. For FACS analysis single cell suspension from bone marrow, spleen and thymus were stained for 20 minutes at 4°C using combinations of antibodies conjugated with FITC, PE, PE-Cy5, Pe-Cy7, APC, APC-Cy7. All samples were pre-incubated for 20 minutes at 4°C with unconjugated CD16/CD32 Fc-blocking antibody to avoid unspecific binding, unless otherwise indicated. The antibodies were purchased from BD bioscience, ebioscience and BioLegend (see Materials). Labelled cells population were discriminated using FACS Canto. After staining cells were washed with FACS buffer, centrifuged at 4°C 10 minutes at 1300 rpm to remove excess of antibody.

For magnetic and non-magnetic cell sorting cells were filtered through a 40 μm filter to avoid stromal cell contamination. Sorted samples were pre-treated with red blood cell lysis buffer (BD Bioscience) or pre-enriched B Cell Isolation Kit, mouse (Miltenyi). Magnetically labelled cells were separated with AutoMACS or QuadroMACS.

### 6.2.3 Red blood cell lysis

Erythrocytes lysis was performed using BD Pharm Lyse<sup>TM</sup> purchased from BD Pharmingen. 10X RBC buffer was diluted using bidistilled water kept at room temperature. 3 ml of 1X

## Materials and methods

RBC lysis buffer were used to treat bone marrow cells derived from one mouse. Bone marrow cells were washed with 1X PBS and centrifuged at RT for 10 minutes at 1300 rpm. Pellets were incubated for 15 minutes at RT with appropriate amount of lysis buffer and then centrifuged RT for 10 minutes at 1300 rpm. Cell pellets were then washed with 1X PBS to remove traces of lysis buffer and broken erythrocytes.

### 6.2.4 Definition of hematopoietic cell types for FACS analysis and FACS sorting

Cell population	Gating strategy
Immature B	*living cells, lymph, doublets, IgD-, B220+, IgM+
Mature B	living cells, lymph, doublets, IgM-, B220+, IgD+
Pro B	living cells, lymph, doublets, CD19+, IgD-, IgM-, c-kit+ CD25-
Pre B	living cells, lymph, doublets, CD19+, IgD-, IgM-, c-kit- CD25+
Fr. A	living cells, lymph, doublets, CD43+, B220+, HSA/CD24-/ low, BP-1-
Fr. B	living cells, lymph, doublets, CD43+, B220+, HSA/CD24+/high, BP-1-
Fr. C	living cells, lymph, doublets, CD43+, B220+, HSA/CD24 high, BP-1+
LSK	living cells, doublets, lin-, Sca+, c-kit+
CLP	living cells, doublets, lin-, IL7 $\alpha$ +, Sca low, c-kit low
CMP	living cells, doublets, lin-, IL7 $\alpha$ -, c-kit high, Sca-, CD34+, CD16/32-/ low
GMP	living cells, doublets, lin-, IL7 $\alpha$ -, c-kit high, Sca-, CD34+, CD16/32+
MEP	living cells, doublets, lin-, IL7 $\alpha$ -, c-kit high, Sca-, CD34-, CD16/32-
Pre-pro	living cells, lymph, doublets, lin-, IL7 $\alpha$ +, c-kit low, CD43low, B220+, CD93+
PF	living cells, lymph, doublets, B220+, PNA+, Fas+
TFH	living cells, lymph, doublets, CD4+, CXCR5+, PD-1+, ICOS+
Naive CD4	living cells, lymph, doublets, CD4+, CD25- , CD62L+, CD44-/low
T mem CD4	living cells, lymph, doublets, CD4+, CD25- , CD62L+, CD44+
T eff like-CD4	living cells, lymph, doublets, CD4+, CD25- , CD62L low/-, CD44+
Naïve CD8	living cells, lymph, doublets, CD8+ , CD62L+, CD44-
T mem CD8	living cells, lymph, doublets, CD8+, CD62L+, CD44+
T eff like-CD8	living cells, lymph, doublets, CD8+, CD62L low/-, CD44+
Immature B sp	living cells, lymph, doublets, CD19, B220, IgM+, IgD-
Mature B sp	living cells, lymph, doublets, CD19, B220, IgM-, IgD+
Trans B sp	living cells, lymph, doublets, CD19, B220, IgM+, IgD low/+
GM	living cells, granulocytes, doublets, Gr+, Mac+
Erythroblasts 1	living cells, c-kit-, Ter119-, CD71-
Erythroblasts 2	living cells, c-kit-, Ter119-, CD71+
Erythroblasts 3	living cells, c-kit-, Ter119+, CD71+
Erythroblasts 4	living cells, c-kit-, Ter119+, CD71-
DN1	living cells, lymph, doublets, CD4-, CD8-, CD44-, CD25-
DN2	living cells, lymph, doublets, CD4-, CD8-, CD44+, CD25-
DN3	living cells, lymph, doublets, CD4-, CD8-, CD44+, CD25+
DN4	living cells, lymph, doublets, CD4-, CD8-, CD44-, CD25+
CD4/CD8	living cells, lymph, doublets, CD4+, CD8+

\*To set the “living cells” gate we performed exclusion of debris and dead cells using FSC and SSC parameters.

### 6.2.5 Envelope protein and Fcγr2b staining

To detect the envelope protein on the pro B cell surface,  $4 \times 10^6$  bone marrow cells were resuspended in a volume of 50  $\mu$ l and incubated 30 minutes at RT with rat  $\alpha$ -env (83A25) diluted 1:5. Subsequently cells were incubated for 1 hour at RT with an anti-rat secondary antibody conjugated Alexa647. Bone marrow cells were then stained with IgD, IgM, CD19 and c-kit to discriminate the pro B cell population (IgD-, IgM-, CD19+ and c-kit+). Because all antibodies used for the pro B cell staining were rat antibodies, env staining had to be performed in the absence of Fc-block and before the pro B cell staining to avoid non-specific bonding of the secondary  $\alpha$ -ratAlexa647.

Detection of the Fcγr2b was also achievable in the absence of Fc-block which directly recognizes and covers Fcγr2b on the cells surface. Pro B cells were stained with the same marker combination described above with the implementation of CD25 to discriminate the pre B cell population (IgD-, IgM-, CD19+, c-kit-, CD25+). After every incubation step, cells were washed with FACS buffer to remove the excess of antibody and centrifuged at 4°C for 10 minutes at 1300 rpm.

### 6.2.6 Bone marrow transplantation

3-4 before transplantation control and mutant mice were injected with 5-fluorouracil (5-FU) to eliminate enriched for low cycling hematopoietic progenitors.

5-FU treated bone marrow cells were harvested from CD45.1 wild type and CD45.2 donor mice. To perform competitive bone marrow transplantation assay,  $1 \times 10^6$  cells from CD45.1 wild type mice were mixed 1:1 with either control (+, +; +/*Mbl*) or *Setdb1*<sup>Mbl</sup> mice. The mixture was transplanted in sublethally irradiated wild type mice through tail vein injection. CD45.1 and CD45.2 surface markers were used to discriminate between the donors. Bone marrow and spleen from recipients were analysed by flow cytometry 7 to 9 weeks after transplantation.

### 6.2.7 Immunohistochemistry

For immunohistochemistry, spleens were fixed overnight with 4% formaldehyde and embedded in paraffin.

To perform haematoxylin/eosin staining embedded sample were cut with a microtome to generate mm section which were deposited onto polarized slide and let adhere. Slides were deparaffinised through rehydration using a descending gradient of ethanol. Splenic sections

were then incubated in haematoxylin solution and washed. Next, slides were counterstained with eosin, washed and dehydrated using an increasing graded series of ethanol. Once dehydration was completed slides were washed in X-TRA-Solv solution and mounted.

To perform ki-67 staining sections were deparaffinised through rehydration using a descending gradient of ethanol and incubated in 10 mM citrate buffer 1 hour at 90°C in the water bath. Samples were then passively cooled at RT. Next, slides were incubated in 3% hydrogen peroxide solution in the dark at RT for 10 minutes. Before proceeding with primary antibody incubation, tissue epitopes were blocked with blocking solution to avoid non-specific binding.  $\alpha$ -ki67 diluted in blocking solution was then distributed on the slides which were incubated overnight at 4°C. The day after samples were washed to remove the excess of antibody and incubated in the dark for 1 hour at RT with secondary antibody conjugated with biotin. A second series of washing step was performed to remove unbound secondary antibody. In the following steps horseradish peroxidase (HRP) conjugated streptavidin was used for detection of the biotinylated secondary antibody. After washing away HRP conjugated streptavidin, slides were incubated with DAB buffer for 10 minutes at RT to allow detection.

#### **6.2.8 Autoantibodies detection test on MEFs**

Blood was quickly taken from the heart of wild type and *Suv420h2 ko* mice. Samples were left at RT to promote coagulation. Afterwards the non-coagulated upper phase (serum) was transferred in a fresh 1.5 ml tube and centrifuged at full speed to eliminate residual traces of cellular components. The upper phase was transferred again in a fresh tube and stored at -20°C. To test the presence of autoantibodies,  $1 \times 10^5$  MEFs were seeded on 12 mm coverslips placed in 24 well plates. Following overnight incubation, medium was aspirate and MEFs were fixed with 3,7% formaldehyde and permeabilized. To avoid non-specific binding fixed cells were incubated for 30 minutes at RT with blocking solution containing bovine serum albumin. 200  $\mu$ l of serum diluted 1:80 from wild type and *Suv420h2 ko* mice were distributed on MEFs and incubated overnight at 4°C. Next day the plate was equilibrated for 30 minutes at RT and incubated with  $\alpha$ -mouse antibody conjugated with Alexa488 for 1 hour. After secondary incubation coverslips were embedded with Vectaschield containing DAPI and sealed with nail polish on a microscope slide.

#### **6.2.9 ELISA**

ELISA test was performed in collaboration with Strobl laboratory (Helmholtz Zentrum, Munich).



### 6.2.10 B cell proliferation assay

MACS depleted B cells were counted and resuspended in appropriate volume to achieve  $20 \times 10^6$  cells/ml. Sterile warm PBS was used to pre-dilute 10 mM CFSE stock solution 1:5000. This solution was then mixed 1:1 ratio with B cell suspension and incubated in the water bath at 37°C 10 minutes. During this step cells incorporated CFSE. Stained cells were then washed with FCS and centrifuged at 4°C 10 minutes at 1300 rpm to stop CFSE labeling. A further washing was performed with supplemented medium. Because CFSE labeling induces cells loss, cell were counted again and suspended at a concentration of  $0,5-1 \times 10^6$  cells/ml in RPMI.  $0,5-1 \times 10^5$  cells were seeded on 96 well plate together with the following immunostimulants: LPS (0,5; 1 and 5 µg/ml); combination of IL-4 (1; 3 and 10 ng/ml) and CD40 (5 µg/ml); IgM (5 µg/ml) and CpG (10 µM). Cell proliferation monitored via CFSE dilution peaks was checked 3-4 days after stimulation with FACS Canto.

### 6.2.11 Colony forming assay and B cell differentiation on OP9 cells

To test B cell differentiation, lineage negative or bulk bone marrow cells treated with red blood cell lysis buffer were seeded ( $1 \times 10^5$  cells/ml) in duplicate in 60 mm dishes on MethoCult 3630 (Stem Cell Technology) containing IL-7 according to the manufacturer's protocol. Cells were grown at 37°C 5% CO<sub>2</sub> and checked every each day to monitor colony formation. After 10-12 day of culture colonies were scored and subsequently analysed by flow cytometry.

Additionally, lineage negative bone marrow cells were enriched via magnetic sorting using the lineage cell depletion kit (Miltenyi) and seeded on OP9 stromal cells in the presence of IL-7 (10 ng/ml). Every other day cells were split in fresh IL-7 supplemented medium as previously described (Holmes and Zuniga-Pflucker, 2009). At day 10 of coculture cells were harvested and analysed by flow cytometry for pre B cell differentiation using Hardy scheme (Hardy et al., 1991).

### 6.2.12 Annexin V staining

$1-2 \times 10^6$  bone marrow cells were pre stained with pro B cell markers and then prepared for Annexin V staining. Next, cells were washed once with PBS and then with 1 ml of 1X Annexin V buffer. Pellets were resuspended in 100 µl of Annexin V buffer to which 5 µl of Annexin V were added. Cells were incubated in the dark at RT for 15 minutes. To remove unbound Annexin V cells were washed with 1-2 ml of Annexin V buffer and immediately analysed by FACS.

## Materials and methods

---

The above described protocol faithfully follows the manufacture's instruction (<http://media.ebioscience.com/data/pdf/best-protocols/annexin-v-staining.pdf>).

### 6.2.13 Cytospin

$1 \times 10^6$  cells were suspended in 200  $\mu$ l of PBS and loaded in a cytospin chamber. Cells were centrifuged at 180 g for 10 minutes. Slides were then stained with May-Grünwald solution for 3 minutes and washed with water. A second staining step was performed with ready to use GIEMSA solution. Excess of the colorant was removed by rinsing with water the slides chamber and samples were air dried before proceeding to microscope analysis.

### 6.2.14 RNA-Seq and qRT-PCR

RNA-Seq was performed in collaboration with Busslinger laboratory (IMP, Vienna). Confirmation of the deregulated genes via qRT-PCR was achieved after RNA purification using RNeasy kit or RNeasy Plus kit from QIAGEN. Both kits allowed elimination of genomic DNA. Purified RNA was then used to produce cDNA. RNA was first incubated at 70°C for 10 minutes with random hexamers. Subsequently, each tube was provided with a mix containing 5X First strand buffer, DTT, dNTPs and RNasin. Half volume from each sample was then transferred in a fresh 1.5 ml tube. One tube was provided with water as negative control while to the other one was given appropriated amount of SuperscriptIII. Samples were incubated for 50 minutes at 50°C and to allow retrotranscription and afterwards 15 minutes at 70°C to stop the reaction. The newly synthesized cDNA was diluted 1:20 to perform qRT-PCR. Quantification of the transcripts was performed using the Ct values obtained using the SYBR green dye.

### 6.2.15 Microarray analysis

Microarray analysis was performed in collaboration with the Patrick Kremer laboratory (Gene Center, Munich).

### 6.2.16 Chromatin immunoprecipitation and ChIP-Seq

ChIP-Seq was performed in collaboration with Busslinger laboratory (IMP, Vienna). For fixation cells were resuspended in PBS containing 1% formaldehyde and 10% FBS. The mixture was incubated 10 minutes at RT with gentle shaking. Fixation reaction was stopped by adding glycine (0.125 M final concentration) for 5 minutes with gentle shaking. Fixed cells were centrifuged 1600 g for 5 minutes and supernatant was discarded. Cell pellets were then washed twice by rotation upon adding PBS 10% FCS. After the last wash, cells were

## Materials and methods

centrifuged again at 1600 g for 5 minutes. Supernatant was carefully removed and pellets were flash frozen.

Before shearing cells were lysed by adding 130  $\mu$ l of buffer B to the pellet and vortexed till complete re-suspension. Sonication was performed using Covaris E220 for 20 minutes. Cell lysates were transferred to AFA Fiber microtubes with Snap-Cap were sheared with the following settings:

Duty cycle	2%
Peak Incident Power	105 Watts (for S220/E220)
Cycles per Burst	200
Processing Time	20 minutes (for $1 \times 10^6$ mouse pro B cells)
Temperature (bath)	4°C
Degassing Mode	Continuous
Volume	130 $\mu$ L in microtubes AFA Fiber with Snap-Cap

For the immunoprecipitation step the antibodies of choice have been bound to magnetic Dynabeads. Next, 100  $\mu$ L of sheared chromatin (diluted in buffer A) were added to the appropriate tubes containing antibody conjugated magnetic beads. Chromatin was incubated for 4 hours at 4°C on a rotator at 35 rpm for immunoprecipitation. All tubes containing immunoprecipitated material were placed in the magnetic rack which captures the beads, to allow proceeding with subsequent washing steps. To remove unwanted immunoprecipitated material, beads were first washed 3 times with 100  $\mu$ l ice cold buffer A and then once with buffer C. After washing, supernatants were discarded and chromatin eluted from the beads by resuspending samples in 210  $\mu$ l of elution buffer and incubating them for 20 minutes at 65°C and 800 rpm in a termomixer. To take out the eluted chromatin, samples were spin down and placed back in the magnetic rack which allows the separation of the beads from the eluted chromatin. The supernatant (200  $\mu$ l) was transferred to a fresh tube and incubated in a PCR machine overnight at 65°C for reverse cross-linking. To digest cellular protein and RNA all samples have been treated with Rnase A and proteinase K. After this step, chromatin has been cleaned using phenol/chloroform purification. 1  $\mu$ l of this material was used for qRT-PCR quantification.

### Buffers composition:

Buffer A - IP and wash buffer (for 10 mL):

- 100  $\mu$ L 1M Tris HCl pH7.5 (10 mM)
- 20  $\mu$ L 0.5M EDTA (1 mM)

## Materials and methods

---

- 10  $\mu$ L 0.5M EGTA (0.5 mM)
- 1000  $\mu$ L 10% Triton 100 X (1%)
- 50  $\mu$ L 20% SDS (0.1%)
- 100  $\mu$ L 0.1% Na-deoxycholate
- 280  $\mu$ L 5M NaCl (140 mM)

Total (1560  $\mu$ L)

- 8240  $\mu$ L H<sub>2</sub>O Milli-Q

\* The 200  $\mu$ L left to complete to 10 mL are completed with protease inhibitor cocktail 50X, but it is added right before the use of the buffer.

Buffer B - lysis buffer (for 10 mL):

- 500  $\mu$ L 1M Tris-HCl pH 8.0 (50 mM)
- 200  $\mu$ L 0.5M EDTA (10 mM)
- 500  $\mu$ L 20% SDS (1%)

Total (1200  $\mu$ L)

- 8600  $\mu$ L H<sub>2</sub>O Milli-Q

The 200  $\mu$ L left to complete to 10 mL are completed with protease inhibitor cocktail 50x, but it is added right before the use of the buffer.

Buffer C - wash buffer (for 10 mL):

- 100  $\mu$ L 1M Tris-HCl pH 8.0 (10 mM)
- 200  $\mu$ L 0.5M EDTA (10 mM)

Total: 1200  $\mu$ L

- 9700  $\mu$ L H<sub>2</sub>O Milli-Q

Elution buffer (for 10 mL):

- 500  $\mu$ L 1M Tris-HCl pH 8.0 (50 mM)
- 200  $\mu$ L 0.5M EDTA (10 mM)
- 500  $\mu$ L 20% SDS (1%)

Total 1200  $\mu$ L

- 8800  $\mu$ L H<sub>2</sub>O Milli-Q

## 7 Abbreviations

5-FU	5-fluorouracil
7AAD	7-Aminoactinomycin
AID	Activation Induced Cytidine Deaminase
AnnV	Annexin V
APC	Antigen presenting cells
APC	Allophycocyanin
APC-Cy7	Allophycocyanin-Cy7 conjugated
<i>Ascl2</i> ; ASCL2	Achaete-scute complex homolog 2
ASCOM	ASC-2/NCOA6 complex
ASH2L	(absent, small, or homeotic)-like
BACH2	BTB and CNC homology 2
BATF	Basic leucine zipper transcription factor, ATF-like
<i>Bcl2</i> ; BCL2	B cell leukemia/lymphoma 2
<i>Bcl6</i> ; BCL6	B cell leukemia/lymphoma 6
BCL-XL	B-cell lymphoma-extra large
BLIMP-1	B-Lymphocyte-Induced Maturation Protein 1
BMI	BMI1 proto-oncogene, polycomb ring finger
C/EBP $\alpha$	CCAAT-enhancer-binding proteins $\alpha$
CCL12	Chemokine (C-C motif) ligand 12
CCR7	Chemokine (C-C motif) receptor 7
cDNA	Complementary DANN
CFSE	Carboxyfluorescein succinimidyl ester
ChIP	Chromatin immunoprecipitation
CLP	Common lymphoid progenitor
c-MAF	Musculoaponeurotic fibrosarcoma AS42 oncogene homolog
CMP	Common myeloid progenitor
CXCL13	Chemokine (C-X-C motif) ligand 13
<i>Cxcr5</i> ; CXCR5	Chemokine (C-X-C motif) receptor 5
DC	Dendritic cell
<i>Def8</i>	Defensin 8
DN	Double negative
DNMT1	DNA methyltransferase (cytosine-5) 1
DNMT2 (TRDMT1)	tRNA aspartic acid methyltransferase 1
DNMT3A	DNA methyltransferase 3A
DNMT3B	DNA methyltransferase 3B
DNMT3L	DNA methyltransferase 3-like
<i>Dntt</i>	Deoxynucleotidyltransferase, terminal
DP	Double positive
E	Embryoni stage
E2A; <i>E2a</i> ( <i>Tcf3</i> )	Transcription factor 3
<i>Ebf1</i> ; EBF1	Early B cell factor 1
EED	Embryonic ectoderm development
EKLF	Erythroid Krüppel-like factor
ELISA	Enzyme-linked immunosorbent assay
ELP	Early lymphoid progenitor
env	Envelope
ERV	Endogenous retroviruses
mESCs	Murine embryonic stem cells
ETS	E26 transformation-specific or E-twenty-six
EZH1/2; <i>Ezh1/2</i>	Enhancer of zeste homolog 1
FACS	Fluorescence activated cell sorting
FACT	Facilitates Chromatin Transcription
FITC	Fluorescein isothiocyanate
FLI-1	Friend leukemia integration 1

## Abbreviations

---

<i>Flt-3</i> ; FLT-3	FMS-related tyrosine kinase 3
<i>Flt-3l</i>	FMS-like tyrosine kinase 3 ligand
FO Z B	Follicular zone B cells
FOG-1	Friend Of GATA-1
FoXO	Forkhead box O
<i>Foxo1</i>	Forkhead box O1
<i>Gata1-3</i> ; GATA1-3	Globin transcription factor 1
<i>Gfi1</i> ; GFI1	Growth factor independent 1
GO	Gene ontology
Gr-1	Granulocyte-1
H	Histone
H3K14ac	Histone 3 Lysine 14 acetylation
H3K18ac	Histone 3 Lysine 18 acetylation
H3K23ac	Histone 3 Lysine 23 acetylation
H3K27me3	Histone 3 Lysine 27 trimethylation
H3K36me3	Histone 3 Lysine 36 trimethylation
H3K4me3	Histone 3 Lysine 4 trimethylation
H3K79me3	Histone 3 Lysine 79 trimethylation
H3K9ac	Histone 3 Lysine 9 acetylation
H3K9me2/3	Histone 3 Lysine 9 dimethylation or trimethylation
H4K12ac	Histone 3 Lysine 12 acetylation
H4K16ac	Histone 3 Lysine 16 acetylation
H4k20me3	Histone 4 Lysine 20 trimethylation
H4K5ac	Histone 4 Lysine 5 acetylation
H4K8ac	Histone 4 Lysine 8 acetylation
HPR	Horseradish peroxidase
HSC	Hematopoietic stem cells
IAP	Internal A-type Particle
IAP-Ez	Internal A-type Particle containing an envelope gene
ICOS	Inducible T cell co-stimulator
<i>Id2/3</i>	Inhibitor of DNA binding 2/3
<i>Ikzf1</i>	IKAROS family zinc finger 1
<i>Ikzf3</i>	IKAROS family zinc finger 3
IL-12	Interleukin 12
IL-21	Interleukin 21
IL-4	Interleukin 4
IL-6	Interleukin 6
<i>Il-7ra</i> ; IL-7R $\alpha$	Interleukin 7 receptor- $\alpha$
Imm B	Immature B cells
INF- $\gamma$	Interferon- $\gamma$
<i>Irf4/8</i> ; IRF4/8	Interferon regulatory factor 4/8
KAP-1 ( <i>Trim28</i> )	KRAB-associated protein-1; tripartite motif-containing 28
Ko	Knockout
KRAB-ZFN	Krüppel associated box-zinc finger
LDB1	LIM domain binding 1
Lin-	Lineage negative
LMPP	Lymphoid mltipotent primed progenitors
LPS	Lipopolysaccharide
LSK	Lineage negative Sca+ c-kit +
LT-HSC	Long term hematopoietic stem cells
LTR	Long terminal repeat
Mac-1	Macrophage-1
MACS	Magnetic activatd cell sorting
Mat B	Mature B cells
MBD	Methyl binding domain
MCL-1	Myeloid cell leukemia sequence 1

## Abbreviations

MEF	Mouse embryonic fibroblast
MegE	Megakaryocyte/erythrocyte
MEP	Megakaryocyte/erythrocyte progenitor
MITF	Microphthalmia-associated transcription factor
<i>Mll</i> ; MLL	Myeloid/lymphoid or mixed-lineage leukemia
MLV or MuLV	Murine leukemia virus
MMTV	Mouse mammary tumor virus
MMVL30	Interspersed repeat subfamily
MPP	Multipotent progenitor
MTA3	Metastasis associated 3
<i>Myb</i>	Myeloblastosis oncogene
<i>Myz</i>	Myc-Interacting Zn Finger Protein-1
MZ B	Marginal zone B cells
NK	Natural killer
NOTCH1-4	Notch Homolog 1-4, Translocation-Associated
<i>Oct25</i>	POU class 5 homeobox 3, gene 2
OP9	Stroma cell line derived from (C57BL/6 x C3H)F2 -op/op mice
PARP- $\gamma$	Poly (ADP-ribose) polymerase- $\gamma$
Pax5; PAX5	Paired box 5
PBS (1)	Phosphate saline buffer
PBS (2)	Primer binding site
PE	Phycoerythrin
PE-Cy5	Phycoerythrin-Cy5 conjugated
PE-Cy7	Phycoerythrin-Cy7 conjugated
PH	Patch deletion region
PRC1	Polycomb-group repressive complex 1
PRC2	Polycomb-group repressive complex 2
Pre B	Precursor B cells
Pre-Pro B	Precursors of progenitor B cells
<i>Prmd1</i>	PR domain containing 1, with ZNF domain
Pro B	Progenitor B cells
PSC	Posterior sex combs
<i>Rag1/2</i> ; RAG1/2	Recombination activating gene 1/2
RbBP5	Retinoblastoma-binding protein 5
RING	Ring finger protein 1
<i>Rorc</i>	RAR-related orphan receptor gamma
RPKM	Reads Per Kilobase of transcript per Million mapped reads
RT	Room temperature
SCML	Sex comb on midleg-like
SET domain	Su(var)3-9 and 'Enhancer of zeste' protein domain
<i>Setdb1</i> ; SETDB1	SET domain, bifurcated 1
<i>Sfp1</i> (PU.1)	Spleen focus forming virus (SFFV) proviral integration oncogene
SLC	Surrogate light chain
SLP65	Src Homology 2 Domain-Containing Leukocyte Protein Of 65 KDa
Sp	Spleen
STAT3	Signal transducer and activator of transcription 3
STAT5 A/B	Signal transducer and activator of transcription 5 A/B
ST-HSC	Short term hematopoietic stem cells
<i>Suv39h</i> ; SUV39H	Suppressor of variegation 3-9 homolog 1
<i>Suv420h1/2</i> ; SUV420H1/2	Suppressor of variegation 4-20 homolog 1/2
SUZ12	Suppressor of zeste 12 homolog
T eff-like	T cell effector-like
T mem	Memory T cell
Tbx21	T-box 21
TEC	Thymic epithelial cell
TF	Transcription factor

## Abbreviations

---

TFH	Follicular helper T cells
TGF $\beta$	Transforming growth factor, beta
TH1	T helper 1
TH17	T helper 17
TH2	T helper 2
TH9	T helper 9
Trans B	Transitory B cells
T reg	Regulatory T cells
TSS	Transcriptional start site
<i>Tubb3</i>	Tubulin, beta 3 class III
URE	Upstream regulatory element
<i>Vav</i>	Vav 1 oncogene
VCAM	Vascular cell adhesion molecule 1
Wnt	Wingless-type
XBP-1	X-box binding protein 1
ZFN	Zinc finger
<i>Zfx</i>	Zinc finger protein X-linked



## 8 Curriculum vitae

The curriculum vitae is not available on the online version.

## 9 Acknowledgments

*“What does not kill you makes you stronger” - Friedrich W. Nietzsche*

This is the famous statement that a German philosopher coined years ago. Many people interpret these words as a mere consolation to bypass depressive states. I do think, instead, that this statement simply depicts life as a constant self-improvement exercise and that the higher is the ambition, the higher is the price we have to pay to become better. Sufferance is simply one of the ways to go to improve ourselves.

During the PhD time sufferance is almost a choice, but at the very end it pays you back. Bad and good times in Gunnar Schotta lab made me what I'm now, and leaving behind disagreements and misunderstanding I'm proud of what I became. That's why I want to express my gratitude to Gunnar for having given me the possibility to go through this experience under his careful supervision.

Next to my supervisor, my colleagues shared with me the daily sacrifices that have to be made on the way to the PhD title. These people believed in what I believed and fought for what I fought, with the same intensity. In particular, it was my pleasure to spend my time with Rui, Gustavo, Dennis and Silvia. In their own way they managed to be part of those memories which will always be able to get a smile out of me. I also thank Mana, Sarah, Filippo, Alex, Maike, Stan and all others old and new members of the Schotta lab for the nice moments spent together and for listening to all my complains. Thanks guys.

A big thanking goes, especially, to my family. They have always believed in me, unconditionally. Thanks to mum and Vittorio to be, always and however, the cornerstones of my existence. Thanks to grandpa Nino e grandma Sandra, that even if they are not there anymore, they continue living in all the good things I make.

I also want to thank all my friends for having being there, next to me, with their minds or also with a scold. Thanks to Viola for having accompanied me and supported me in this adventure beyond the Alps. Thanks to Sara, Alessio, Mara, Paolino, Susy, Dario e Alessandro for having represented the best selection of Italians that, in whatever country, I would like to take with me.

Thanks to Gerti and Toni Harasim, for having welcoming me with open arms beyond the border.

Thanks to Thomas, for the priceless support that can only be provided by certain kind of feelings. This work is entirely dedicated to him.

Last but not least, I thank myself.....I thank myself for never giving up and because I totally believe that many people could have done better with my projects, but only few of them would have made it with the same determination, passion and dedication.

*“Genius is 1% inspiration and 99% perspiration” - Thomas A. Edison*

## 10 Citations

Adolfsson, J., Borge, O.J., Bryder, D., Theilgaard-Monch, K., Astrand-Grundstrom, I., Sitnicka, E., Sasaki, Y., and Jacobsen, S.E. (2001). Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity* 15, 659-669.

Adolfsson, J., Mansson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.J., Thoren, L.A., *et al.* (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell* 121, 295-306.

Akashi, K., Traver, D., Miyamoto, T., and Weissman, I.L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* 404, 193-197.

Allan, R.S., Zueva, E., Cammas, F., Schreiber, H.A., Masson, V., Belz, G.T., Roche, D., Maison, C., Quivy, J.P., Almouzni, G., *et al.* (2012). An epigenetic silencing pathway controlling T helper 2 cell lineage commitment. *Nature* 487, 249-253.

Allman, D., Sambandam, A., Kim, S., Miller, J.P., Pagan, A., Well, D., Meraz, A., and Bhandoola, A. (2003). Thymopoiesis independent of common lymphoid progenitors. *Nature immunology* 4, 168-174.

Ang, Y.S., Tsai, S.Y., Lee, D.F., Monk, J., Su, J., Ratnakumar, K., Ding, J., Ge, Y., Darr, H., Chang, B., *et al.* (2011). Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* 145, 183-197.

Back, J., Allman, D., Chan, S., and Kastner, P. (2005). Visualizing PU.1 activity during hematopoiesis. *Experimental hematology* 33, 395-402.

Bain, G., Maandag, E.C., Izon, D.J., Amsen, D., Kruisbeek, A.M., Weintraub, B.C., Krop, I., Schlissel, M.S., Feeney, A.J., van Roon, M., *et al.* (1994). E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. *Cell* 79, 885-892.

Balciunaite, G., Ceredig, R., Massa, S., and Rolink, A.G. (2005). A B220+ CD117+ CD19-hematopoietic progenitor with potent lymphoid and myeloid developmental potential. *European journal of immunology* 35, 2019-2030.

Bauquet, A.T., Jin, H., Paterson, A.M., Mitsdoerffer, M., Ho, I.C., Sharpe, A.H., and Kuchroo, V.K. (2009). The costimulatory molecule ICOS regulates the expression of c-Maf and IL-21 in the development of follicular T helper cells and TH-17 cells. *Nature immunology* 10, 167-175.

Beckmann, J., Scheitza, S., Wernet, P., Fischer, J.C., and Giebel, B. (2007). Asymmetric cell division within the human hematopoietic stem and progenitor cell compartment: identification of asymmetrically segregating proteins. *Blood* 109, 5494-5501.

Beguelin, W., Popovic, R., Teater, M., Jiang, Y., Bunting, K.L., Rosen, M., Shen, H., Yang, S.N., Wang, L., Ezponda, T., *et al.* (2013). EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer cell* 23, 677-692.

Belotserkovskaya, R., and Reinberg, D. (2004). Facts about FACT and transcript elongation through chromatin. *Current opinion in genetics & development* 14, 139-146.

- Berger, S.L., Kouzarides, T., Shiekhata, R., and Shilatifard, A. (2009). An operational definition of epigenetics. *Genes & development* 23, 781-783.
- Bestor, T.H. (2000). The DNA methyltransferases of mammals. *Human molecular genetics* 9, 2395-2402.
- Betz, B.C., Jordan-Williams, K.L., Wang, C., Kang, S.G., Liao, J., Logan, M.R., Kim, C.H., and Taparowsky, E.J. (2010). Batf coordinates multiple aspects of B and T cell function required for normal antibody responses. *The Journal of experimental medicine* 207, 933-942.
- Bhandoola, A., von Boehmer, H., Petrie, H.T., and Zuniga-Pflucker, J.C. (2007). Commitment and developmental potential of extrathymic and intrathymic T cell precursors: plenty to choose from. *Immunity* 26, 678-689.
- Biard-Piechaczyk, M., Robert-Hebmann, V., Richard, V., Roland, J., Hipskind, R.A., and Devaux, C. (2000). Caspase-dependent apoptosis of cells expressing the chemokine receptor CXCR4 is induced by cell membrane-associated human immunodeficiency virus type 1 envelope glycoprotein (gp120). *Virology* 268, 329-344.
- Bilodeau, S., Kagey, M.H., Frampton, G.M., Rahl, P.B., and Young, R.A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes & development* 23, 2484-2489.
- Bollig, N., Brustle, A., Kellner, K., Ackermann, W., Abass, E., Raifer, H., Camara, B., Brendel, C., Giel, G., Bothur, E., *et al.* (2012). Transcription factor IRF4 determines germinal center formation through follicular T-helper cell differentiation. *Proceedings of the National Academy of Sciences of the United States of America* 109, 8664-8669.
- Bostick, M., Kim, J.K., Esteve, P.O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317, 1760-1764.
- Bracken, A.P., and Helin, K. (2009). Polycomb group proteins: navigators of lineage pathways led astray in cancer. *Nature reviews Cancer* 9, 773-784.
- Broske, A.-M., Vockentanz, L., Kharazi, S., Huska, M.R., Mancini, E., Scheller, M., Kuhl, C., Enns, A., Prinz, M., Jaenisch, R., *et al.* (2009). DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nature genetics* 41, 1207-1215.
- Brummendorf, T.H., Dragowska, W., Zijlmans, J., Thornbury, G., and Lansdorp, P.M. (1998). Asymmetric cell divisions sustain long-term hematopoiesis from single-sorted human fetal liver cells. *The Journal of experimental medicine* 188, 1117-1124.
- Cairns, B.R. (2009). The logic of chromatin architecture and remodelling at promoters. *Nature* 461, 193-198.
- Carpenter, A.C., and Bosselut, R. (2010). Decision checkpoints in the thymus. *Nature immunology* 11, 666-673.
- Cattoretti, G., Chang, C.C., Cechova, K., Zhang, J., Ye, B.H., Falini, B., Louie, D.C., Offit, K., Chaganti, R.S., and Dalla-Favera, R. (1995). BCL-6 protein is expressed in germinal-center B cells. *Blood* 86, 45-53.

## Citations

---

Celli, S., Lemaitre, F., and Bousso, P. (2007). Real-time manipulation of T cell-dendritic cell interactions in vivo reveals the importance of prolonged contacts for CD4<sup>+</sup> T cell activation. *Immunity* 27, 625-634.

Chang, A.N., Cantor, A.B., Fujiwara, Y., Lodish, M.B., Droho, S., Crispino, J.D., and Orkin, S.H. (2002). GATA-factor dependence of the multitype zinc-finger protein FOG-1 for its essential role in megakaryopoiesis. *Proceedings of the National Academy of Sciences of the United States of America* 99, 9237-9242.

Chen, C., Nott, T.J., Jin, J., and Pawson, T. (2011). Deciphering arginine methylation: Tudor tells the tale. *Nature reviews Molecular cell biology* 12, 629-642.

Chuang, L.S., Ian, H.I., Koh, T.W., Ng, H.H., Xu, G., and Li, B.F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* 277, 1996-2000.

Clark, M.R., Mandal, M., Ochiai, K., and Singh, H. (2014). Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. *Nature reviews Immunology* 14, 69-80.

Clements, W.K., and Traver, D. (2013). Signalling pathways that control vertebrate haematopoietic stem cell specification. *Nature reviews Immunology* 13, 336-348.

Cobaleda, C., Schebesta, A., Delogu, A., and Busslinger, M. (2007). Pax5: the guardian of B cell identity and function. *Nature immunology* 8, 463-470.

Crotty, S. (2011). Follicular helper CD4 T cells (TFH). *Annual review of immunology* 29, 621-663.

Dakic, A., Metcalf, D., Di Rago, L., Mifsud, S., Wu, L., and Nutt, S.L. (2005). PU.1 regulates the commitment of adult hematopoietic progenitors and restricts granulopoiesis. *The Journal of experimental medicine* 201, 1487-1502.

Dambacher, S., Hahn, M., and Schotta, G. (2010). Epigenetic regulation of development by histone lysine methylation. *Heredity* 105, 24-37.

Dambacher, S., Hahn, M., and Schotta, G. (2013). The compact view on heterochromatin. *Cell cycle* 12, 2925-2926.

DeKoter, R.P., and Singh, H. (2000). Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science* 288, 1439-1441.

Dent, A.L., Shaffer, A.L., Yu, X., Allman, D., and Staudt, L.M. (1997). Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science* 276, 589-592.

Dias, S., Silva, H., Jr., Cumano, A., and Vieira, P. (2005). Interleukin-7 is necessary to maintain the B cell potential in common lymphoid progenitors. *The Journal of experimental medicine* 201, 971-979.

Ding, L., and Morrison, S.J. (2013). Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature* 495, 231-235.

- Dodge, J.E., Kang, Y.K., Beppu, H., Lei, H., and Li, E. (2004). Histone H3-K9 methyltransferase ESET is essential for early development. *Mol Cell Biol* 24, 2478-2486.
- Dou, Y., Milne, T.A., Ruthenburg, A.J., Lee, S., Lee, J.W., Verdine, G.L., Allis, C.D., and Roeder, R.G. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nature structural & molecular biology* 13, 713-719.
- Dranoff, G. (2004). Cytokines in cancer pathogenesis and cancer therapy. *Nature reviews Cancer* 4, 11-22.
- Dupont, C., Armant, D.R., and Brenner, C.A. (2009). Epigenetics: definition, mechanisms and clinical perspective. *Seminars in reproductive medicine* 27, 351-357.
- Egle, A., Harris, A.W., Bath, M.L., O'Reilly, L., and Cory, S. (2004). VavP-Bcl2 transgenic mice develop follicular lymphoma preceded by germinal center hyperplasia. *Blood* 103, 2276-2283.
- Ellyard, J.I., and Vinuesa, C.G. (2011). A BATF-ling connection between B cells and follicular helper T cells. *Nature immunology* 12, 519-520.
- Eto, D., Lao, C., DiToro, D., Barnett, B., Escobar, T.C., Kageyama, R., Yusuf, I., and Crotty, S. (2011). IL-21 and IL-6 are critical for different aspects of B cell immunity and redundantly induce optimal follicular helper CD4 T cell (Tfh) differentiation. *PloS one* 6, e17739.
- Ezhkova, E., Pasolli, H.A., Parker, J.S., Stokes, N., Su, I.H., Hannon, G., Tarakhovsky, A., and Fuchs, E. (2009). Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells. *Cell* 136, 1122-1135.
- Fahl, S.P., Crittenden, R.B., Allman, D., and Bender, T.P. (2009). c-Myb is required for pro-B cell differentiation. *Journal of immunology (Baltimore, Md : 1950)* 183, 5582-5592.
- Faust, C., Schumacher, A., Holdener, B., and Magnuson, T. (1995). The eed mutation disrupts anterior mesoderm production in mice. *Development* 121, 273-285.
- Ficara, F., Murphy, M.J., Lin, M., and Cleary, M.L. (2008). Pbx1 regulates self-renewal of long-term hematopoietic stem cells by maintaining their quiescence. *Cell stem cell* 2, 484-496.
- Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes & development* 17, 1870-1881.
- Friedman, A.D. (2002). Transcriptional regulation of granulocyte and monocyte development. *Oncogene* 21, 3377-3390.
- Friedman, J.R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G., and Rauscher, F.J., 3rd (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes & development* 10, 2067-2078.
- Fujita, N., Jaye, D.L., Geigerman, C., Akyildiz, A., Mooney, M.R., Boss, J.M., and Wade, P.A. (2004). MTA3 and the Mi-2/NuRD complex regulate cell fate during B lymphocyte differentiation. *Cell* 119, 75-86.
- Fujiwara, Y., Browne, C.P., Cunniff, K., Goff, S.C., and Orkin, S.H. (1996). Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor

GATA-1. Proceedings of the National Academy of Sciences of the United States of America 93, 12355-12358.

Fukuda, T., Yoshida, T., Okada, S., Hatano, M., Miki, T., Ishibashi, K., Okabe, S., Koseki, H., Hirosawa, S., Taniguchi, M., *et al.* (1997). Disruption of the Bcl6 gene results in an impaired germinal center formation. The Journal of experimental medicine 186, 439-448.

Fuxa, M., Skok, J., Souabni, A., Salvagiotto, G., Roldan, E., and Busslinger, M. (2004). Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. Genes & development 18, 411-422.

Galan-Caridad, J.M., Harel, S., Arenzana, T.L., Hou, Z.E., Doetsch, F.K., Mirny, L.A., and Reizis, B. (2007). Zfx controls the self-renewal of embryonic and hematopoietic stem cells. Cell 129, 345-357.

Georgiades, P., Ogilvy, S., Duval, H., Licence, D.R., Charnock-Jones, D.S., Smith, S.K., and Print, C.G. (2002). VavCre transgenic mice: a tool for mutagenesis in hematopoietic and endothelial lineages. Genesis (New York, NY : 2000) 34, 251-256.

Georgopoulos, K., Bigby, M., Wang, J.H., Molnar, A., Wu, P., Winandy, S., and Sharpe, A. (1994). The Ikaros gene is required for the development of all lymphoid lineages. Cell 79, 143-156.

Georgopoulos, K., Moore, D.D., and Derfler, B. (1992). Ikaros, an early lymphoid-specific transcription factor and a putative mediator for T cell commitment. Science 258, 808-812.

Georgopoulos, K., Winandy, S., and Avitahl, N. (1997). The role of the Ikaros gene in lymphocyte development and homeostasis. Annual review of immunology 15, 155-176.

Glaser, S., Schaft, J., Lubitz, S., Vintersten, K., van der Hoeven, F., Tufteland, K.R., Aasland, R., Anastassiadis, K., Ang, S.L., and Stewart, A.F. (2006). Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. Development 133, 1423-1432.

Glass, C.K., and Rosenfeld, M.G. (2000). The coregulator exchange in transcriptional functions of nuclear receptors. Genes & development 14, 121-141.

Goldberg, A.D., Allis, C.D., and Bernstein, E. (2007). Epigenetics: a landscape takes shape. Cell 128, 635-638.

Gonda, H., Sugai, M., Nambu, Y., Katakai, T., Agata, Y., Mori, K.J., Yokota, Y., and Shimizu, A. (2003). The balance between Pax5 and Id2 activities is the key to AID gene expression. The Journal of experimental medicine 198, 1427-1437.

Goyama, S., Yamamoto, G., Shimabe, M., Sato, T., Ichikawa, M., Ogawa, S., Chiba, S., and Kurokawa, M. (2008). Evi-1 is a critical regulator for hematopoietic stem cells and transformed leukemic cells. Cell stem cell 3, 207-220.

Graf, T., McNagny, K., Brady, G., and Frampton, J. (1992). Chicken "erythroid" cells transformed by the Gag-Myb-Ets-encoding E26 leukemia virus are multipotent. Cell 70, 201-213.

Groner, A.C., Meylan, S., Ciuffi, A., Zangger, N., Ambrosini, G., Denervaud, N., Bucher, P., and Trono, D. (2010). KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. PLoS genetics 6, e1000869.

- Gyory, I., Boller, S., Nechanitzky, R., Mandel, E., Pott, S., Liu, E., and Grosschedl, R. (2012). Transcription factor Ebf1 regulates differentiation stage-specific signaling, proliferation, and survival of B cells. *Genes & development* 26, 668-682.
- Hahn, M., Dambacher, S., Dulev, S., Kuznetsova, A.Y., Eck, S., Worz, S., Sadic, D., Schulte, M., Mallm, J.P., Maiser, A., *et al.* (2013). Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin. *Genes & development* 27, 859-872.
- Hardtke, S., Ohl, L., and Forster, R. (2005). Balanced expression of CXCR5 and CCR7 on follicular T helper cells determines their transient positioning to lymph node follicles and is essential for efficient B-cell help. *Blood* 106, 1924-1931.
- Hardy, R.R., Carmack, C.E., Shinton, S.A., Kemp, J.D., and Hayakawa, K. (1991). Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *The Journal of experimental medicine* 173, 1213-1225.
- Harigae, H. (2006). GATA transcription factors and hematological diseases. *The Tohoku journal of experimental medicine* 210, 1-9.
- Harman, B.C., Miller, J.P., Nikbakht, N., Gerstein, R., and Allman, D. (2006). Mouse plasmacytoid dendritic cells derive exclusively from estrogen-resistant myeloid progenitors. *Blood* 108, 878-885.
- Harris, M.B., Chang, C.C., Berton, M.T., Danial, N.N., Zhang, J., Kuehner, D., Ye, B.H., Kvatyuk, M., Pandolfi, P.P., Cattoretti, G., *et al.* (1999). Transcriptional repression of Stat6-dependent interleukin-4-induced genes by BCL-6: specific regulation of iepsilon transcription and immunoglobulin E switching. *Mol Cell Biol* 19, 7264-7275.
- Hayday, A.C., and Pennington, D.J. (2007). Key factors in the organized chaos of early T cell development. *Nature immunology* 8, 137-144.
- Heesters, B.A., Myers, R.C., and Carroll, M.C. (2014). Follicular dendritic cells: dynamic antigen libraries. *Nature reviews Immunology* 14, 495-504.
- Heltemes-Harris, L.M., Willette, M.J., Vang, K.B., and Farrar, M.A. (2011). The role of STAT5 in the development, function, and transformation of B and T lymphocytes. *Annals of the New York Academy of Sciences* 1217, 18-31.
- Herquel, B., Ouarrhni, K., Martianov, I., Le Gras, S., Ye, T., Keime, C., Lerouge, T., Jost, B., Cammas, F., Losson, R., *et al.* (2013). Trim24-repressed VL30 retrotransposons regulate gene expression by producing noncoding RNA. *Nature structural & molecular biology* 20, 339-346.
- Herzog, S., Hug, E., Meixlsperger, S., Paik, J.H., DePinho, R.A., Reth, M., and Jumaa, H. (2008). SLP-65 regulates immunoglobulin light chain gene recombination through the PI(3)K-PKB-Foxo pathway. *Nature immunology* 9, 623-631.
- Herzog, S., Reth, M., and Jumaa, H. (2009). Regulation of B-cell proliferation and differentiation by pre-B-cell receptor signalling. *Nature reviews Immunology* 9, 195-205.
- Heuser, M., Yap, D.B., Leung, M., de Algora, T.R., Tafech, A., McKinney, S., Dixon, J., Thresher, R., Colledge, B., Carlton, M., *et al.* (2009). Loss of MLL5 results in pleiotropic hematopoietic defects, reduced neutrophil immune function, and extreme sensitivity to DNA demethylation. *Blood* 113, 1432-1443.



## Citations

---

- Hirokawa, S., Sato, H., Kato, I., and Kudo, A. (2003). EBF-regulating Pax5 transcription is enhanced by STAT5 in the early stage of B cells. *European journal of immunology* 33, 1824-1829.
- Hobeika, E., Thiemann, S., Storch, B., Jumaa, H., Nielsen, P.J., Pelanda, R., and Reth, M. (2006). Testing gene function early in the B cell lineage in mb1-cre mice. *Proceedings of the National Academy of Sciences of the United States of America* 103, 13789-13794.
- Hock, H., Hamblen, M.J., Rooke, H.M., Schindler, J.W., Saleque, S., Fujiwara, Y., and Orkin, S.H. (2004). Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. *Nature* 431, 1002-1007.
- Hock, H., Hamblen, M.J., Rooke, H.M., Traver, D., Bronson, R.T., Cameron, S., and Orkin, S.H. (2003). Intrinsic requirement for zinc finger transcription factor Gfi-1 in neutrophil differentiation. *Immunity* 18, 109-120.
- Holmes, M.L., Carotta, S., Corcoran, L.M., and Nutt, S.L. (2006). Repression of Flt3 by Pax5 is crucial for B-cell lineage commitment. *Genes & development* 20, 933-938.
- Holmes, R., and Zuniga-Pflucker, J.C. (2009). The OP9-DL1 system: generation of T-lymphocytes from embryonic or hematopoietic stem cells in vitro. *Cold Spring Harbor protocols* 2009, pdb.prot5156.
- Horcher, M., Souabni, A., and Busslinger, M. (2001). Pax5/BSAP maintains the identity of B cells in late B lymphopoiesis. *Immunity* 14, 779-790.
- Huang, C., Geng, H., Boss, I., Wang, L., and Melnick, A. (2014). Cooperative transcriptional repression by BCL6 and BACH2 in germinal center B-cell differentiation. *Blood* 123, 1012-1020.
- Igarashi, H., Medina, K.L., Yokota, T., Rossi, M.I., Sakaguchi, N., Comp, P.C., and Kincade, P.W. (2005). Early lymphoid progenitors in mouse and man are highly sensitive to glucocorticoids. *International immunology* 17, 501-511.
- Iscove, N.N., and Nawa, K. (1997). Hematopoietic stem cells expand during serial transplantation in vivo without apparent exhaustion. *Current biology : CB* 7, 805-808.
- Iwasaki, H., and Akashi, K. (2007). Myeloid lineage commitment from the hematopoietic stem cell. *Immunity* 26, 726-740.
- Iwasaki, H., Somoza, C., Shigematsu, H., Duprez, E.A., Iwasaki-Arai, J., Mizuno, S., Arinobu, Y., Geary, K., Zhang, P., Dayaram, T., *et al.* (2005). Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood* 106, 1590-1600.
- Iyengar, S., Ivanov, A.V., Jin, V.X., Rauscher, F.J., 3rd, and Farnham, P.J. (2011). Functional analysis of KAP1 genomic recruitment. *Mol Cell Biol* 31, 1833-1847.
- Janeway, C. (2008a). The Development and Survival of Lymphocytes. In *Immunobiology : the immune system in health and disease* (United States: Garland science).
- Janeway, C. (2008b). *Immunobiology : the immune system in health and disease*, 5 edn (United States: Garland science).

Janeway, C.A., Jr. (2001). How the immune system works to protect the host from infection: a personal view. *Proceedings of the National Academy of Sciences of the United States of America* 98, 7461-7468.

Jelicic, K., Cimbro, R., Nawaz, F., Huang, D.W., Zheng, X., Yang, J., Lempicki, R.A., Pascuccio, M., Van Ryk, D., Schwing, C., *et al.* (2013). The HIV-1 envelope protein gp120 impairs B cell proliferation by inducing TGF- $\beta$ 1 production and FcRL4 expression. *Nature immunology* 14, 1256-1265.

Jenkinson, E.J., Jenkinson, W.E., Rossi, S.W., and Anderson, G. (2006). The thymus and T-cell commitment: the right niche for Notch? *Nature reviews Immunology* 6, 551-555.

Jiang, C., and Pugh, B.F. (2009). Nucleosome positioning and gene regulation: advances through genomics. *Nature reviews Genetics* 10, 161-172.

Johnston, R.J., Poholek, A.C., DiToro, D., Yusuf, I., Eto, D., Barnett, B., Dent, A.L., Craft, J., and Crotty, S. (2009). Bcl6 and Blimp-1 are reciprocal and antagonistic regulators of T follicular helper cell differentiation. *Science* 325, 1006-1010.

Joshi, I., Yoshida, T., Jena, N., Qi, X., Zhang, J., Van Etten, R.A., and Georgopoulos, K. (2014). Loss of Ikaros DNA-binding function confers integrin-dependent survival on pre-B cells and progression to acute lymphoblastic leukemia. *Nature immunology* 15, 294-304.

Jude, C.D., Climer, L., Xu, D., Artinger, E., Fisher, J.K., and Ernst, P. (2007). Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. *Cell stem cell* 1, 324-337.

Kallies, A., Hasbold, J., Tarlinton, D.M., Dietrich, W., Corcoran, L.M., Hodgkin, P.D., and Nutt, S.L. (2004). Plasma cell ontogeny defined by quantitative changes in blimp-1 expression. *The Journal of experimental medicine* 200, 967-977.

Kang, Y.K. (2014). SETDB1 in Early Embryos and Embryonic Stem Cells. *Current issues in molecular biology* 17, 1-10.

Karimi, M.M., Goyal, P., Maksakova, I.A., Bilenky, M., Leung, D., Tang, J.X., Shinkai, Y., Mager, D.L., Jones, S., Hirst, M., *et al.* (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell stem cell* 8, 676-687.

Kass, S.U., Pruss, D., and Wolffe, A.P. (1997). How does DNA methylation repress transcription? *Trends in genetics* : TIG 13, 444-449.

Katada, S., and Sassone-Corsi, P. (2010). The histone methyltransferase MLL1 permits the oscillation of circadian gene expression. *Nature structural & molecular biology* 17, 1414-1421.

Kee, B.L. (2009). E and ID proteins branch out. *Nature reviews Immunology* 9, 175-184.

Kikuchi, K., Lai, A.Y., Hsu, C.L., and Kondo, M. (2005). IL-7 receptor signaling is necessary for stage transition in adult B cell development through up-regulation of EBF. *The Journal of experimental medicine* 201, 1197-1203.

Kim, J., and Kim, H. (2012). Recruitment and biological consequences of histone modification of H3K27me3 and H3K9me3. *ILAR journal / National Research Council, Institute of Laboratory Animal Resources* 53, 232-239.

## Citations

---

Kinoshita, K., and Honjo, T. (2001). Linking class-switch recombination with somatic hypermutation. *Nature reviews Molecular cell biology* 2, 493-503.

Kirstetter, P., Thomas, M., Dierich, A., Kastner, P., and Chan, S. (2002). Ikaros is critical for B cell differentiation and function. *European journal of immunology* 32, 720-730.

Klein, U., and Dalla-Favera, R. (2008). Germinal centres: role in B-cell physiology and malignancy. *Nature reviews Immunology* 8, 22-33.

Kondo, M., Weissman, I.L., and Akashi, K. (1997). Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell* 91, 661-672.

Kosan, C., Saba, I., Godmann, M., Herold, S., Herkert, B., Eilers, M., and Moroy, T. (2010). Transcription factor miz-1 is required to regulate interleukin-7 receptor signaling at early commitment stages of B cell differentiation. *Immunity* 33, 917-928.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.

Krivtsov, A.V., and Armstrong, S.A. (2007). MLL translocations, histone modifications and leukaemia stem-cell development. *Nature reviews Cancer* 7, 823-833.

Kueh, H.Y., Champhekar, A., Nutt, S.L., Elowitz, M.B., and Rothenberg, E.V. (2013). Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. *Science* 341, 670-673.

Kulesa, H., Frampton, J., and Graf, T. (1995). GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblats, and erythroblats. *Genes & development* 9, 1250-1262.

Kwon, H., Thierry-Mieg, D., Thierry-Mieg, J., Kim, H.P., Oh, J., Tunyaplin, C., Carotta, S., Donovan, C.E., Goldman, M.L., Tailor, P., *et al.* (2009). Analysis of interleukin-21-induced *Prdm1* gene regulation reveals functional cooperation of STAT3 and IRF4 transcription factors. *Immunity* 31, 941-952.

Lai, A.Y., and Kondo, M. (2006). Asymmetrical lymphoid and myeloid lineage commitment in multipotent hematopoietic progenitors. *The Journal of experimental medicine* 203, 1867-1873.

Lai, A.Y., and Kondo, M. (2007). Identification of a bone marrow precursor of the earliest thymocytes in adult mouse. *Proceedings of the National Academy of Sciences of the United States of America* 104, 6311-6316.

Laiosa, C.V., Stadtfeld, M., and Graf, T. (2006). Determinants of lymphoid-myeloid lineage diversification. *Annual review of immunology* 24, 705-738.

Lawrenz-Smith, S.C., and Thomas, C.Y. (1995). The E47 transcription factor binds to the enhancer sequences of recombinant murine leukemia viruses and influences enhancer function. *Journal of virology* 69, 4142-4148.

Lawson, K.A., Teteak, C.J., Gao, J., Li, N., Hacquebord, J., Ghatan, A., Zielinska-Kwiatkowska, A., Song, G., Chansky, H.A., and Yang, L. (2013a). ESET histone methyltransferase regulates osteoblastic differentiation of mesenchymal stem cells during postnatal bone development. *FEBS letters* 587, 3961-3967.

Lawson, K.A., Teteak, C.J., Zou, J., Hacquebord, J., Ghatan, A., Zielinska-Kwiatkowska, A., Fernandes, R.J., Chansky, H.A., and Yang, L. (2013b). Mesenchyme-specific knockout of ESET histone methyltransferase causes ectopic hypertrophy and terminal differentiation of articular chondrocytes. *The Journal of biological chemistry* 288, 32119-32125.

Lee, J., Saha, P.K., Yang, Q.H., Lee, S., Park, J.Y., Suh, Y., Lee, S.K., Chan, L., Roeder, R.G., and Lee, J.W. (2008). Targeted inactivation of MLL3 histone H3-Lys-4 methyltransferase activity in the mouse reveals vital roles for MLL3 in adipogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 105, 19229-19234.

Lemischka, I.R., Raulet, D.H., and Mulligan, R.C. (1986). Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell* 45, 917-927.

Lewis, E.B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* 276, 565-570.

Li, J. (2011). Quiescence regulators for hematopoietic stem cell. *Experimental hematology* 39, 511-520.

Li, L., Jothi, R., Cui, K., Lee, J.Y., Cohen, T., Gorivodsky, M., Tzchori, I., Zhao, Y., Hayes, S.M., Bresnick, E.H., *et al.* (2011). Nuclear adaptor Ldb1 regulates a transcriptional program essential for the maintenance of hematopoietic stem cells. *Nature immunology* 12, 129-136.

Li, P., Spolski, R., Liao, W., Wang, L., Murphy, T.L., Murphy, K.M., and Leonard, W.J. (2012). BATF-JUN is critical for IRF4-mediated transcription in T cells. *Nature* 490, 543-546.

Li, Y.S., Wasserman, R., Hayakawa, K., and Hardy, R.R. (1996). Identification of the earliest B lineage stage in mouse bone marrow. *Immunity* 5, 527-535.

Lin, H., and Grosschedl, R. (1995). Failure of B-cell differentiation in mice lacking the transcription factor EBF. *Nature* 376, 263-267.

Lin, L., Gerth, A.J., and Peng, S.L. (2004). Active inhibition of plasma cell development in resting B cells by microphthalmia-associated transcription factor. *The Journal of experimental medicine* 200, 115-122.

Lin, Y.C., Jhunjhunwala, S., Benner, C., Heinz, S., Welinder, E., Mansson, R., Sigvardsson, M., Hagman, J., Espinoza, C.A., Dutkowski, J., *et al.* (2010). A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nature immunology* 11, 635-643.

Liu, H., Westergard, T.D., and Hsieh, J.J. (2009). MLL5 governs hematopoiesis: a step closer. *Blood* 113, 1395-1396.

Liu, X., Chen, X., Zhong, B., Wang, A., Wang, X., Chu, F., Nurieva, R.I., Yan, X., Chen, P., van der Flier, L.G., *et al.* (2014). Transcription factor achaete-scute homologue 2 initiates follicular T-helper-cell development. *Nature* 507, 513-518.

Liu, X., Nurieva, R.I., and Dong, C. (2013). Transcriptional regulation of follicular T-helper (Tfh) cells. *Immunological reviews* 252, 139-145.

Lu, K.T., Kanno, Y., Cannons, J.L., Handon, R., Bible, P., Elkahouloun, A.G., Anderson, S.M., Wei, L., Sun, H., O'Shea, J.J., *et al.* (2011). Functional and epigenetic studies reveal multistep

## Citations

---

differentiation and plasticity of in vitro-generated and in vivo-derived follicular T helper cells. *Immunity* 35, 622-632.

Lu, R., Medina, K.L., Lancki, D.W., and Singh, H. (2003). IRF-4,8 orchestrate the pre-B-to-B transition in lymphocyte development. *Genes & development* 17, 1703-1708.

Lu, X., Sachs, F., Ramsay, L., Jacques, P.-É., Göke, J., Bourque, G., and Ng, H.-H. (2014). The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. *Nature structural & molecular biology* 21, 423-425.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.

Ma, C.S., Avery, D.T., Chan, A., Batten, M., Bustamante, J., Boisson-Dupuis, S., Arkwright, P.D., Kreins, A.Y., Averbuch, D., Engelhard, D., *et al.* (2012a). Functional STAT3 deficiency compromises the generation of human T follicular helper cells. *Blood* 119, 3997-4008.

Ma, C.S., Deenick, E.K., Batten, M., and Tangye, S.G. (2012b). The origins, function, and regulation of T follicular helper cells. *The Journal of experimental medicine* 209, 1241-1253.

Ma, S., Pathak, S., Trinh, L., and Lu, R. (2008). Interferon regulatory factors 4 and 8 induce the expression of Ikaros and Aiolos to down-regulate pre-B-cell receptor and promote cell-cycle withdrawal in pre-B-cell development. *Blood* 111, 1396-1403.

Ma, S., Turetsky, A., Trinh, L., and Lu, R. (2006). IFN regulatory factor 4 and 8 promote Ig light chain kappa locus activation in pre-B cell development. *Journal of immunology* (Baltimore, Md : 1950) 177, 7898-7904.

Mackarechtschian, K., Hardin, J.D., Moore, K.A., Boast, S., Goff, S.P., and Lemischka, I.R. (1995). Targeted disruption of the flk2/flt3 gene leads to deficiencies in primitive hematopoietic progenitors. *Immunity* 3, 147-161.

MacLennan, I.C., Toellner, K.M., Cunningham, A.F., Serre, K., Sze, D.M., Zuniga, E., Cook, M.C., and Vinuesa, C.G. (2003). Extrafollicular antibody responses. *Immunological reviews* 194, 8-18.

Madan, V., Madan, B., Brykczynska, U., Zilbermann, F., Hogeveen, K., Dohner, K., Dohner, H., Weber, O., Blum, C., Rodewald, H.R., *et al.* (2009). Impaired function of primitive hematopoietic cells in mice lacking the Mixed-Lineage-Leukemia homolog MLL5. *Blood* 113, 1444-1454.

Mak, K.S., Funnell, A.P., Pearson, R.C., and Crossley, M. (2011). PU.1 and Haematopoietic Cell Fate: Dosage Matters. *International journal of cell biology* 2011, 808524.

Malin, S., McManus, S., and Busslinger, M. (2010). STAT5 in B cell development and leukemia. *Current opinion in immunology* 22, 168-176.

Mansson, R., Hultquist, A., Luc, S., Yang, L., Anderson, K., Kharazi, S., Al-Hashmi, S., Liuba, K., Thoren, L., Adolfsson, J., *et al.* (2007). Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity* 26, 407-419.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. *Nature* 469, 343-349.

- McEwen, K.R., and Ferguson-Smith, A.C. (2010). Distinguishing epigenetic marks of developmental and imprinting regulation. *Epigenetics & chromatin* 3, 2.
- McHeyzer-Williams, M., Okitsu, S., Wang, N., and McHeyzer-Williams, L. (2012). Molecular programming of B cell memory. *Nature reviews Immunology* 12, 24-34.
- McKenna, H.J., Stocking, K.L., Miller, R.E., Brasel, K., De Smedt, T., Maraskovsky, E., Maliszewski, C.R., Lynch, D.H., Smith, J., Pulendran, B., *et al.* (2000). Mice lacking flt3 ligand have deficient hematopoiesis affecting hematopoietic progenitor cells, dendritic cells, and natural killer cells. *Blood* 95, 3489-3497.
- McKercher, S.R., Torbett, B.E., Anderson, K.L., Henkel, G.W., Vestal, D.J., Baribault, H., Klemsz, M., Feeney, A.J., Wu, G.E., Paige, C.J., *et al.* (1996). Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *The EMBO journal* 15, 5647-5658.
- McMahon, K.A., Hiew, S.Y., Hadjur, S., Veiga-Fernandes, H., Menzel, U., Price, A.J., Kioussis, D., Williams, O., and Brady, H.J. (2007). Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. *Cell stem cell* 1, 338-345.
- Mebius, R.E., and Kraal, G. (2005). Structure and function of the spleen. *Nature reviews Immunology* 5, 606-616.
- Medina, K.L., Garrett, K.P., Thompson, L.F., Rossi, M.I., Payne, K.J., and Kincade, P.W. (2001). Identification of very early lymphoid precursors in bone marrow and their regulation by estrogen. *Nature immunology* 2, 718-724.
- Medvedovic, J., Ebert, A., Tagoh, H., and Busslinger, M. (2011). Pax5: a master regulator of B cell development and leukemogenesis. *Advances in immunology* 111, 179-206.
- Melchers, F. (2005). The pre-B-cell receptor: selector of fitting immunoglobulin heavy chains for the B-cell repertoire. *Nature reviews Immunology* 5, 578-584.
- Mikkola, H.K., and Orkin, S.H. (2006). The journey of developing hematopoietic stem cells. *Development* 133, 3733-3744.
- Miller, J.F. (1961). Immunological function of the thymus. *Lancet* 2, 748-749.
- Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D., and Hess, J.L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. *Molecular cell* 10, 1107-1117.
- Min, J., Zhang, Y., and Xu, R.M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes & development* 17, 1823-1828.
- Moazed, D., and O'Farrell, P.H. (1992). Maintenance of the engrailed expression pattern by Polycomb group genes in *Drosophila*. *Development* 116, 805-810.
- Molnar, A., and Georgopoulos, K. (1994). The Ikaros gene encodes a family of functionally diverse zinc finger DNA-binding proteins. *Mol Cell Biol* 14, 8292-8303.
- Mombaerts, P., Iacomini, J., Johnson, R.S., Herrup, K., Tonegawa, S., and Papaioannou, V.E. (1992). RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* 68, 869-877.

## Citations

---

Morrison, S.J., and Scadden, D.T. (2014). The bone marrow niche for haematopoietic stem cells. *Nature* 505, 327-334.

Mueller, S.N., and Germain, R.N. (2009). Stromal cell contributions to the homeostasis and functionality of the immune system. *Nature reviews Immunology* 9, 618-629.

Muto, A., Tashiro, S., Nakajima, O., Hoshino, H., Takahashi, S., Sakoda, E., Ikebe, D., Yamamoto, M., and Igarashi, K. (2004). The transcriptional programme of antibody class switching involves the repressor Bach2. *Nature* 429, 566-571.

Nanua, S., and Yoshimura, F.K. (2004). Mink epithelial cell killing by pathogenic murine leukemia viruses involves endoplasmic reticulum stress. *Journal of virology* 78, 12071-12074.

Nei, Y., Obata-Ninomiya, K., Tsutsui, H., Ishiwata, K., Miyasaka, M., Matsumoto, K., Nakae, S., Kanuka, H., Inase, N., and Karasuyama, H. (2013). GATA-1 regulates the generation and function of basophils. *Proceedings of the National Academy of Sciences of the United States of America* 110, 18620-18625.

Nerlov, C., and Graf, T. (1998). PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors. *Genes & development* 12, 2403-2412.

Nerlov, C., Querfurth, E., Kulesa, H., and Graf, T. (2000). GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. *Blood* 95, 2543-2551.

Newell-Price, J., Clark, A.J., and King, P. (2000). DNA methylation and silencing of gene expression. *Trends in endocrinology and metabolism: TEM* 11, 142-148.

Nicetto, D., Hahn, M., Jung, J., Schneider, T.D., Straub, T., David, R., Schotta, G., and Rupp, R.A. (2013). Suv4-20h histone methyltransferases promote neuroectodermal differentiation by silencing the pluripotency-associated Oct-25 gene. *PLoS genetics* 9, e1003188.

Nurieva, R.I., Chung, Y., Hwang, D., Yang, X.O., Kang, H.S., Ma, L., Wang, Y.H., Watowich, S.S., Jetten, A.M., Tian, Q., *et al.* (2008). Generation of T follicular helper cells is mediated by interleukin-21 but independent of T helper 1, 2, or 17 cell lineages. *Immunity* 29, 138-149.

Nurieva, R.I., Chung, Y., Martinez, G.J., Yang, X.O., Tanaka, S., Matskevitch, T.D., Wang, Y.H., and Dong, C. (2009). Bcl6 mediates the development of T follicular helper cells. *Science* 325, 1001-1005.

Nutt, S.L., Eberhard, D., Horcher, M., Rolink, A.G., and Busslinger, M. (2001). Pax5 determines the identity of B cells from the beginning to the end of B-lymphopoiesis. *International reviews of immunology* 20, 65-82.

Nutt, S.L., Heavey, B., Rolink, A.G., and Busslinger, M. (1999). Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature*.

Nutt, S.L., and Kee, B.L. (2007). The transcriptional regulation of B cell lineage commitment. *Immunity* 26, 715-725.

Nutt, S.L., Metcalf, D., D'Amico, A., Polli, M., and Wu, L. (2005). Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. *The Journal of experimental medicine* 201, 221-231.

Nutt, S.L., Morrison, A.M., Dorfler, P., Rolink, A., and Busslinger, M. (1998). Identification of BSAP (Pax-5) target genes in early B-cell development by loss- and gain-of-function experiments. *The EMBO journal* *17*, 2319-2333.

Nutt, S.L., and Tarlinton, D.M. (2011). Germinal center B and follicular helper T cells: siblings, cousins or just good friends? *Nature immunology* *13*, 472-477.

Nutt, S.L., Urbanek, P., Rolink, A., and Busslinger, M. (1997). Essential functions of Pax5 (BSAP) in pro-B cell development: difference between fetal and adult B lymphopoiesis and reduced V-to-DJ recombination at the IgH locus. *Genes & development* *11*, 476-491.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The polycomb-group gene *Ezh2* is required for early mouse development. *Mol Cell Biol* *21*, 4330-4336.

O'Carroll, D., Scherthan, H., Peters, A.H., Opravil, S., Haynes, A.R., Laible, G., Rea, S., Schmid, M., Lebersorger, A., Jerratsch, M., *et al.* (2000). Isolation and characterization of Suv39h2, a second histone H3 methyltransferase gene that displays testis-specific expression. *Mol Cell Biol* *20*, 9423-9433.

O'Riordan, M., and Grosschedl, R. (1999). Coordinate regulation of B cell differentiation by the transcription factors EBF and E2A. *Immunity* *11*, 21-31.

Ochiai, K., Maienschein-Cline, M., Simonetti, G., Chen, J., Rosenthal, R., Brink, R., Chong, A.S., Klein, U., Dinner, A.R., Singh, H., *et al.* (2013). Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. *Immunity* *38*, 918-929.

Odegard, J.M., Marks, B.R., DiPlacido, L.D., Poholek, A.C., Kono, D.H., Dong, C., Flavell, R.A., and Craft, J. (2008). ICOS-dependent extrafollicular helper T cells elicit IgG production via IL-21 in systemic autoimmunity. *The Journal of experimental medicine* *205*, 2873-2886.

Ogilvy, S., Metcalf, D., Print, C.G., Bath, M.L., Harris, A.W., and Adams, J.M. (1999). Constitutive Bcl-2 expression throughout the hematopoietic compartment affects multiple lineages and enhances progenitor cell survival. *Proceedings of the National Academy of Sciences of the United States of America* *96*, 14943-14948.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* *99*, 247-257.

Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* *132*, 631-644.

Park, I.K., Qian, D., Kiel, M., Becker, M.W., Pihalja, M., Weissman, I.L., Morrison, S.J., and Clarke, M.F. (2003). Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* *423*, 302-305.

Parker, M.J., Licence, S., Erlandsson, L., Galler, G.R., Chakalova, L., Osborne, C.S., Morgan, G., Fraser, P., Jumaa, H., Winkler, T.H., *et al.* (2005). The pre-B-cell receptor induces silencing of VpreB and lambda5 transcription. *The EMBO journal* *24*, 3895-3905.

Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *The EMBO journal* *23*, 4061-4071.



## Citations

---

Pelegri, F., and Lehmann, R. (1994). A role of polycomb group genes in the regulation of gap gene expression in *Drosophila*. *Genetics* 136, 1341-1353.

Perfettini, J.L., Castedo, M., Roumier, T., Andreau, K., Nardacci, R., Piacentini, M., and Kroemer, G. (2005). Mechanisms of apoptosis induction by the HIV-1 envelope. *Cell death and differentiation* 12 Suppl 1, 916-923.

Pevny, L., Simon, M.C., Robertson, E., Klein, W.H., Tsai, S.F., D'Agati, V., Orkin, S.H., and Costantini, F. (1991). Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* 349, 257-260.

Pichlmair, A., and Reis e Sousa, C. (2007). Innate recognition of viruses. *Immunity* 27, 370-383.

Pietras, E.M., Warr, M.R., and Passegue, E. (2011). Cell cycle regulation in hematopoietic stem cells. *The Journal of cell biology* 195, 709-720.

Polli, M., Dakic, A., Light, A., Wu, L., Tarlinton, D.M., and Nutt, S.L. (2005). The development of functional B lymphocytes in conditional PU.1 knock-out mice. *Blood* 106, 2083-2090.

Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. *Nature biotechnology* 28, 1057-1068.

Pui, J.C., Allman, D., Xu, L., DeRocco, S., Karnell, F.G., Bakkour, S., Lee, J.Y., Kadesch, T., Hardy, R.R., Aster, J.C., *et al.* (1999). Notch1 expression in early lymphopoiesis influences B versus T lineage determination. *Immunity* 11, 299-308.

Radbruch, A., Muehlinghaus, G., Luger, E.O., Inamine, A., Smith, K.G., Dorner, T., and Hiepe, F. (2006). Competence and competition: the challenge of becoming a long-lived plasma cell. *Nature reviews Immunology* 6, 741-750.

Radtke, F., Wilson, A., Mancini, S.J., and MacDonald, H.R. (2004). Notch regulation of lymphocyte development and function. *Nature immunology* 5, 247-253.

Radtke, F., Wilson, A., Stark, G., Bauer, M., van Meerwijk, J., MacDonald, H.R., and Aguet, M. (1999). Deficient T cell fate specification in mice with an induced inactivation of Notch1. *Immunity* 10, 547-558.

Rampalli, S., Li, L., Mak, E., Ge, K., Brand, M., Tapscott, S.J., and Dilworth, F.J. (2007). p38 MAPK signaling regulates recruitment of Ash2L-containing methyltransferase complexes to specific genes during differentiation. *Nature structural & molecular biology* 14, 1150-1156.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., *et al.* (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593-599.

Reimold, A.M., Ponath, P.D., Li, Y.S., Hardy, R.R., David, C.S., Strominger, J.L., and Glimcher, L.H. (1996). Transcription factor B cell lineage-specific activator protein regulates the gene for human X-box binding protein 1. *The Journal of experimental medicine* 183, 393-401.

- Rekhtman, N., Radparvar, F., Evans, T., and Skoultschi, A.I. (1999). Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. *Genes & development* 13, 1398-1411.
- Reljic, R., Wagner, S.D., Peakman, L.J., and Fearon, D.T. (2000). Suppression of signal transducer and activator of transcription 3-dependent B lymphocyte terminal differentiation by BCL-6. *The Journal of experimental medicine* 192, 1841-1848.
- Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annual review of genetics* 38, 413-443.
- Robertson, K.D. (2005). DNA methylation and human disease. *Nature reviews Genetics* 6, 597-610.
- Roessler, S., Gyory, I., Imhof, S., Spivakov, M., Williams, R.R., Busslinger, M., Fisher, A.G., and Grosschedl, R. (2007). Distinct promoters mediate the regulation of Ebf1 gene expression by interleukin-7 and Pax5. *Mol Cell Biol* 27, 579-594.
- Rolink, A.G., Nutt, S.L., Melchers, F., and Busslinger, M. (1999). Long-term in vivo reconstitution of T-cell development by Pax5-deficient B-cell progenitors. *Nature* 401, 603-606.
- Rosenbauer, F., Owens, B.M., Yu, L., Tumang, J.R., Steidl, U., Kutok, J.L., Clayton, L.K., Wagner, K., Scheller, M., Iwasaki, H., *et al.* (2006). Lymphoid cell growth and transformation are suppressed by a key regulatory element of the gene encoding PU.1. *Nature genetics* 38, 27-37.
- Rosenbauer, F., and Tenen, D.G. (2007). Transcription factors in myeloid development: balancing differentiation with transformation. *Nature reviews Immunology* 7, 105-117.
- Rosenfeld, M.G., Lunyak, V.V., and Glass, C.K. (2006). Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes & development* 20, 1405-1428.
- Rothenberg, E.V. (2007). Cell lineage regulators in B and T cell development. *Nature immunology* 8, 441-444.
- Rothenberg, E.V. (2010). B cell specification from the genome up. *Nature immunology* 11, 572-574.
- Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P.V., Layard-Liesching, H., Verp, S., Marquis, J., *et al.* (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* 463, 237-240.
- Rowe, H.M., and Trono, D. (2011). Dynamic control of endogenous retroviruses during development. *Virology* 411, 273-287.
- Rumfelt, L.L., Zhou, Y., Rowley, B.M., Shinton, S.A., and Hardy, R.R. (2006). Lineage specification and plasticity in CD19- early B cell precursors. *The Journal of experimental medicine* 203, 675-687.
- Ryan, F.P. (2004). Human endogenous retroviruses in health and disease: a symbiotic perspective. *Journal of the Royal Society of Medicine* 97, 560-565.

## Citations

---

Schatz, D.G., and Ji, Y. (2011). Recombination centres and the orchestration of V(D)J recombination. *Nature reviews Immunology* 11, 251-263.

Scheller, M., Huelsken, J., Rosenbauer, F., Taketo, M.M., Birchmeier, W., Tenen, D.G., and Leutz, A. (2006). Hematopoietic stem cell and multilineage defects generated by constitutive beta-catenin activation. *Nature immunology* 7, 1037-1047.

Schlenner, S.M., and Rodewald, H.R. (2010). Early T cell development and the pitfalls of potential. *Trends in immunology* 31, 303-310.

Schlissel, M.S. (2003). Regulating antigen-receptor gene assembly. *Nature reviews Immunology* 3, 890-899.

Schmitt, N., Bentebibel, S.E., and Ueno, H. (2014). Phenotype and functions of memory Tfh cells in human blood. *Trends in immunology* 35, 436-442.

Schones, D.E., and Zhao, K. (2008). Genome-wide approaches to studying chromatin modifications. *Nature reviews Genetics* 9, 179-191.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes & development* 18, 1251-1262.

Schotta, G., Sengupta, R., Kubicek, S., Malin, S., Kauer, M., Callen, E., Celeste, A., Pagani, M., Opravil, S., De La Rosa-Velazquez, I.A., *et al.* (2008). A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes & development* 22, 2048-2061.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes & development* 16, 919-932.

Sciammas, R., and Davis, M.M. (2004). Modular nature of Blimp-1 in the regulation of gene expression during B cell maturation. *Journal of immunology (Baltimore, Md : 1950)* 172, 5427-5440.

Scott, E.W., Simon, M.C., Anastasi, J., and Singh, H. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* 265, 1573-1577.

Seet, C.S., Brumbaugh, R.L., and Kee, B.L. (2004). Early B cell factor promotes B lymphopoiesis with reduced interleukin 7 responsiveness in the absence of E2A. *The Journal of experimental medicine* 199, 1689-1700.

Segal, E., and Widom, J. (2009). What controls nucleosome positions? *Trends in genetics : TIG* 25, 335-343.

Semerad, C.L., Mercer, E.M., Inlay, M.A., Weissman, I.L., and Murre, C. (2009). E2A proteins maintain the hematopoietic stem cell pool and promote the maturation of myelolymphoid and myeloerythroid progenitors. *Proceedings of the National Academy of Sciences of the United States of America* 106, 1930-1935.

Sethi, A., Kulkarni, N., Sonar, S., and Lal, G. (2013). Role of miRNAs in CD4 T cell plasticity during inflammation and tolerance. *Frontiers in genetics* 4.

Shaffer, A.L., Shapiro-Shelef, M., Iwakoshi, N.N., Lee, A.H., Qian, S.B., Zhao, H., Yu, X., Yang, L., Tan, B.K., Rosenwald, A., *et al.* (2004). XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity* 21, 81-93.

Shaffer, A.L., Yu, X., He, Y., Boldrick, J., Chan, E.P., and Staudt, L.M. (2000). BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. *Immunity* 13, 199-212.

Shapiro-Shelef, M., and Calame, K. (2005). Regulation of plasma-cell development. *Nature reviews Immunology* 5, 230-242.

Shapiro-Shelef, M., Lin, K.I., McHeyzer-Williams, L.J., Liao, J., McHeyzer-Williams, M.G., and Calame, K. (2003). Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells. *Immunity* 19, 607-620.

Shen, X., Liu, Y., Hsu, Y.J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.C., and Orkin, S.H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Molecular cell* 32, 491-502.

Shinkai, Y., Rathbun, G., Lam, K.P., Oltz, E.M., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., Stall, A.M., *et al.* (1992). RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* 68, 855-867.

Shivdasani, R.A., Fujiwara, Y., McDevitt, M.A., and Orkin, S.H. (1997). A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *The EMBO journal* 16, 3965-3973.

Singh, H., Medina, K.L., and Pongubala, J.M. (2005). Contingent gene regulatory networks and B cell fate specification. *Proceedings of the National Academy of Sciences of the United States of America* 102, 4949-4953.

Smith, E., and Shilatifard, A. (2010). The chromatin signaling pathway: diverse mechanisms of recruitment of histone-modifying enzymes and varied biological outcomes. *Molecular cell* 40, 689-701.

Sripathy, S.P., Stevens, J., and Schultz, D.C. (2006). The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. *Mol Cell Biol* 26, 8623-8638.

Starck, J., Cohet, N., Gonnet, C., Sarrazin, S., Doubeikovskaia, Z., Doubeikovski, A., Verger, A., Duterque-Coquillaud, M., and Morle, F. (2003). Functional Cross-Antagonism between Transcription Factors FLI-1 and EKLF. *Molecular and Cellular Biology* 23, 1390-1402.

Stehling-Sun, S., Dade, J., Nutt, S.L., DeKoter, R.P., and Camargo, F.D. (2009). Regulation of lymphoid versus myeloid fate 'choice' by the transcription factor Mef2c. *Nature immunology* 10, 289-296.

Stender, J.D., Pascual, G., Liu, W., Kaikkonen, M.U., Do, K., Spann, N.J., Boutros, M., Perrimon, N., Rosenfeld, M.G., and Glass, C.K. (2012). Control of proinflammatory gene programs by regulated trimethylation and demethylation of histone H4K20. *Molecular cell* 48, 28-38.

## Citations

---

Stetson, D.B., and Medzhitov, R. (2006). Type I interferons in host defense. *Immunity* 25, 373-381.

Stopka, T., Amanatullah, D.F., Papetti, M., and Skoultschi, A.I. (2005). PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure. *The EMBO journal* 24, 3712-3723.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41-45.

Straussman, R., Nejman, D., Roberts, D., Steinfeld, I., Blum, B., Benvenisty, N., Simon, I., Yakhini, Z., and Cedar, H. (2009). Developmental programming of CpG island methylation profiles in the human genome. *Nature structural & molecular biology* 16, 564-571.

Struhl, K., and Segal, E. (2013). Determinants of nucleosome positioning. *Nature structural & molecular biology* 20, 267-273.

Subtil-Rodriguez, A., and Reyes, J.C. (2010). BRG1 helps RNA polymerase II to overcome a nucleosomal barrier during elongation, in vivo. *EMBO reports* 11, 751-757.

Swain, S.L., McKinstry, K.K., and Strutt, T.M. (2012). Expanding roles for CD4(+) T cells in immunity to viruses. *Nature reviews Immunology* 12, 136-148.

Takahama, Y. (2006). Journey through the thymus: stromal guides for T-cell development and selection. *Nature reviews Immunology* 6, 127-135.

Takahashi, S., Shimizu, R., Suwabe, N., Kuroha, T., Yoh, K., Ohta, J., Nishimura, S., Lim, K.C., Engel, J.D., and Yamamoto, M. (2000). GATA factor transgenes under GATA-1 locus control rescue germline GATA-1 mutant deficiencies. *Blood* 96, 910-916.

Takano, H., Ema, H., Sudo, K., and Nakauchi, H. (2004). Asymmetric division and lineage commitment at the level of hematopoietic stem cells: inference from differentiation in daughter cell and granddaughter cell pairs. *The Journal of experimental medicine* 199, 295-302.

Tan, S.L., Nishi, M., Ohtsuka, T., Matsui, T., Takemoto, K., Kamio-Miura, A., Aburatani, H., Shinkai, Y., and Kageyama, R. (2012). Essential roles of the histone methyltransferase ESET in the epigenetic control of neural progenitor cells during development. *Development* 139, 3806-3816.

Terranova, R., Agherbi, H., Boned, A., Meresse, S., and Djabali, M. (2006). Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll. *Proceedings of the National Academy of Sciences of the United States of America* 103, 6629-6634.

Thal, M.A., Carvalho, T.L., He, T., Kim, H.G., Gao, H., Hagman, J., and Klug, C.A. (2009). Ebf1-mediated down-regulation of Id2 and Id3 is essential for specification of the B cell lineage. *Proceedings of the National Academy of Sciences of the United States of America* 106, 552-557.

Thomas, M.D., Kremer, C.S., Ravichandran, K.S., Rajewsky, K., and Bender, T.P. (2005). c-Myb is critical for B cell development and maintenance of follicular B cells. *Immunity* 23, 275-286.

Thompson, E.C., Cobb, B.S., Sabbattini, P., Meixlsperger, S., Parelho, V., Liberg, D., Taylor, B., Dillon, N., Georgopoulos, K., Jumaa, H., *et al.* (2007). Ikaros DNA-binding proteins as integral components of B cell developmental-stage-specific regulatory circuits. *Immunity* 26, 335-344.

Tothova, Z., Kollipara, R., Huntly, B.J., Lee, B.H., Castrillon, D.H., Cullen, D.E., McDowell, E.P., Lazo-Kallanian, S., Williams, I.R., Sears, C., *et al.* (2007). FoxOs are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress. *Cell* 128, 325-339.

Traver, D., and Akashi, K. (2004). Lineage commitment and developmental plasticity in early lymphoid progenitor subsets. *Advances in immunology* 83, 1-54.

Tsang, A.P., Fujiwara, Y., Hom, D.B., and Orkin, S.H. (1998). Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. *Genes & development* 12, 1176-1188.

Tsang, A.P., Visvader, J.E., Turner, C.A., Fujiwara, Y., Yu, C., Weiss, M.J., Crossley, M., and Orkin, S.H. (1997). FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell* 90, 109-119.

Tunayaplin, C., Shaffer, A.L., Angelin-Duclos, C.D., Yu, X., Staudt, L.M., and Calame, K.L. (2004). Direct repression of *prdm1* by Bcl-6 inhibits plasmacytic differentiation. *Journal of immunology* (Baltimore, Md : 1950) 173, 1158-1165.

Turner, C.A., Jr., Mack, D.H., and Davis, M.M. (1994). Blimp-1, a novel zinc finger-containing protein that can drive the maturation of B lymphocytes into immunoglobulin-secreting cells. *Cell* 77, 297-306.

Ullrich, C.K., Groopman, J.E., and Ganju, R.K. (2000). HIV-1 gp120- and gp160-induced apoptosis in cultured endothelial cells is mediated by caspases. *Blood* 96, 1438-1442.

Urbanek, P., Wang, Z.Q., Fetka, I., Wagner, E.F., and Busslinger, M. (1994). Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell* 79, 901-912.

van de Lagemaat, L.N., Landry, J.R., Mager, D.L., and Medstrand, P. (2003). Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends in genetics : TIG* 19, 530-536.

Vasanwala, F.H., Kusam, S., Toney, L.M., and Dent, A.L. (2002). Repression of AP-1 function: a mechanism for the regulation of Blimp-1 expression and B lymphocyte differentiation by the B cell lymphoma-6 protooncogene. *Journal of immunology* (Baltimore, Md : 1950) 169, 1922-1929.

Vilagos, B., Hoffmann, M., Souabni, A., Sun, Q., Werner, B., Medvedovic, J., Bilic, I., Minnich, M., Axelsson, E., Jaritz, M., *et al.* (2012). Essential role of EBF1 in the generation and function of distinct mature B cell types. *The Journal of experimental medicine* 209, 775-792.

Vinuesa, C.G., Cook, M.C., Angelucci, C., Athanasopoulos, V., Rui, L., Hill, K.M., Yu, D., Domaschensz, H., Whittle, B., Lambe, T., *et al.* (2005). A RING-type ubiquitin ligase family member required to repress follicular helper T cells and autoimmunity. *Nature* 435, 452-458.

## Citations

---

Vinuesa, C.G., and Cyster, J.G. (2011). How T cells earn the follicular rite of passage. *Immunity* 35, 671-680.

Vinuesa, C.G., Linterman, M.A., Goodnow, C.C., and Randall, K.L. (2010). T cells and follicular dendritic cells in germinal center B-cell formation and selection. *Immunological reviews* 237, 72-89.

Voigt, P., Tee, W.W., and Reinberg, D. (2013). A double take on bivalent promoters. *Genes & development* 27, 1318-1338.

von Andrian, U.H., and Mempel, T.R. (2003). Homing and cellular traffic in lymph nodes. *Nature reviews Immunology* 3, 867-878.

Waddington, C.H. (1942). The epigenotype. *Endeavour* 1, 18-20.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* 431, 873-878.

Wang, J.H., Nichogiannopoulou, A., Wu, L., Sun, L., Sharpe, A.H., Bigby, M., and Georgopoulos, K. (1996). Selective defects in the development of the fetal and adult lymphoid system in mice with an Ikaros null mutation. *Immunity* 5, 537-549.

Wang, N.D., Finegold, M.J., Bradley, A., Ou, C.N., Abdelsayed, S.V., Wilde, M.D., Taylor, L.R., Wilson, D.R., and Darlington, G.J. (1995). Impaired energy homeostasis in C/EBP alpha knockout mice. *Science* 269, 1108-1112.

Welner, R.S., Pelayo, R., and Kincade, P.W. (2008). Evolving views on the genealogy of B cells. *Nature reviews Immunology* 8, 95-106.

Wilson, A., Laurenti, E., Oser, G., van der Wath, R.C., Blanco-Bose, W., Jaworski, M., Offner, S., Dunant, C.F., Eshkind, L., Bockamp, E., *et al.* (2008). Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* 135, 1118-1129.

Wolf, D., and Goff, S.P. (2009). Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature* 458, 1201-1204.

Xu, J., and Orkin, S.H. (2011). The erythroid/myeloid lineage fate paradigm takes a new player. *The EMBO journal* 30, 983-985.

Yang, L., Bryder, D., Adolfsson, J., Nygren, J., Mansson, R., Sigvardsson, M., and Jacobsen, S.E. (2005). Identification of Lin(-)Sca1(+)kit(+)CD34(+)Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. *Blood* 105, 2717-2723.

Ye, M., Ermakova, O., and Graf, T. (2005). PU.1 is not strictly required for B cell development and its absence induces a B-2 to B-1 cell switch. *The Journal of experimental medicine* 202, 1411-1422.

Yoshida, T., Ng, S.Y., Zuniga-Pflucker, J.C., and Georgopoulos, K. (2006). Early hematopoietic lineage restrictions directed by Ikaros. *Nature immunology* 7, 382-391.

Yoshimura, F.K., Wang, T., and Nanua, S. (2001). Mink cell focus-forming murine leukemia virus killing of mink cells involves apoptosis and superinfection. *Journal of virology* 75, 6007-6015.

Young, G.R., Eksmond, U., Salcedo, R., Alexopoulou, L., Stoye, J.P., and Kassiotis, G. (2012). Resurrection of endogenous retroviruses in antibody-deficient mice. *Nature* 491, 774-778.

Yu, B.D., Hanson, R.D., Hess, J.L., Horning, S.E., and Korsmeyer, S.J. (1998). MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 95, 10632-10636.

Yu, B.D., Hess, J.L., Horning, S.E., Brown, G.A., and Korsmeyer, S.J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* 378, 505-508.

Yu, D., Rao, S., Tsai, L.M., Lee, S.K., He, Y., Sutcliffe, E.L., Srivastava, M., Linterman, M., Zheng, L., Simpson, N., *et al.* (2009). The transcriptional repressor Bcl-6 directs T follicular helper cell lineage commitment. *Immunity* 31, 457-468.

Zandi, S., Mansson, R., Tsapogas, P., Zetterblad, J., Bryder, D., and Sigvardsson, M. (2008). EBF1 is essential for B-lineage priming and establishment of a transcription factor network in common lymphoid progenitors. *Journal of immunology (Baltimore, Md : 1950)* 181, 3364-3372.

Zeng, H., Yucel, R., Kosan, C., Klein-Hitpass, L., and Moroy, T. (2004). Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. *The EMBO journal* 23, 4116-4125.

Zhang, D.E., Zhang, P., Wang, N.D., Hetherington, C.J., Darlington, G.J., and Tenen, D.G. (1997). Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice. *Proceedings of the National Academy of Sciences of the United States of America* 94, 569-574.

Zhang, P., Behre, G., Pan, J., Iwama, A., Wara-Aswapati, N., Radomska, H.S., Auron, P.E., Tenen, D.G., and Sun, Z. (1999). Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. *Proceedings of the National Academy of Sciences of the United States of America* 96, 8705-8710.

Zhang, P., Iwasaki-Arai, J., Iwasaki, H., Fenyus, M.L., Dayaram, T., Owens, B.M., Shigematsu, H., Levantini, E., Huettner, C.S., Lekstrom-Himes, J.A., *et al.* (2004). Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. *Immunity* 21, 853-863.

Zhang, P., Zhang, X., Iwama, A., Yu, C., Smith, K.A., Mueller, B.U., Narravula, S., Torbett, B.E., Orkin, S.H., and Tenen, D.G. (2000). PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding. *Blood* 96, 2641-2648.

Zhang, Y., Wong, J., Klinger, M., Tran, M.T., Shannon, K.M., and Killeen, N. (2009). MLL5 contributes to hematopoietic stem cell fitness and homeostasis. *Blood* 113, 1455-1463.

Zhao, X., and Yoshimura, F.K. (2008). Expression of murine leukemia virus envelope protein is sufficient for the induction of apoptosis. *Journal of virology* 82, 2586-2589.



## Citations

---

Zhou, L., Chong, M.M., and Littman, D.R. (2009). Plasticity of CD4<sup>+</sup> T cell lineage differentiation. *Immunity* 30, 646-655.